Essays on Investment Behavior in Online Social Platforms

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVESITY OF MINNESOTA
BY

Swanand J. Deodhar

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Mani Subramani

April 2018

Acknowledgments

I dedicate this dissertation to my beloved parents Shalini J. Deodhar and Janardan N. Deodhar, as well as to my much better half Shaveta Sawhney. Their relentless push and support are the reasons as to why I undertook this journey.

I wish to express my sincerest gratitude to my advisor Mani Subramani. He is not only an extremely knowledgeable professor but also a very caring mentor. His keen insights and drive to push the work forward, his timely words of compassion, and his willingness to regularly dedicate a significant amount of time have been invaluable, to say the least. I also want to thank Aks Zaheer for sparing his precious time and providing highly intellectual advice throughout my Ph.D. journey. He has significantly contributed in helping me craft insightful research questions. Indeed, it has been a real privilege to be mentored by with Aks. I also want thank Yuqing Ren and Jeylan Mortimer for their timely advice and help. They have played an extremely crucial role in getting the dissertation to where it has reached.

I also want to acknowledge the incredible support that the other faculty members in the Information and Decision Sciences Department have extended from time to time. Their advice has played a vital role in all aspects of my academic journey at Carlson. Next, I want to especially thank Kate, Ashley, and Angie for their prompt administrative support. Their tireless effort ensured that my Ph.D. journey went along without any administrative hiccups. Lastly, I want to thank my Ph.D. colleagues Yash, Yaqiong, Sandeep, Scott, Cathy, Ali, Zach, Mochen, Probal, and Ke. I will always cherish the bonds we have formed here.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1 Users' Susceptibility to Others' Actions: A Literature Classification Framework

1. INTRODUCTION

Interpersonal interactions have been the cornerstone of research in a variety of disciplines, including psychology, sociology, economics, and management. Naturally, any fundamental shift in this regard can have far-reaching research and practical implications. In the last decade or so, one such shift has been the emergence of online platforms. By building on the advances in the internet technologies, online platforms have emerged as the new routes for interactions between different social entities, ranging from individuals to governments. For instance, Facebook and MySpace allow individuals to form online social networks, share personal information through various forms of content, and create groups and communities. LinkedIn and GlassDoor extend the same functionalities to professional networks through which individuals can seek employment, know more about a potential employer, and review their current and past employers. Twitter allows users to share news at a speed that, at times, outperforms the traditional news media.

The considerable variation in the activities, which online platforms facilitate is matched by the array of research questions that different literature streams have examined. For social psychologists, the central research question is how different stimuli ingrained in most human interactions play out in online platforms. As McFarland and Ployhart (2015) argue, this research question has a potential to not only theoretically distinguish online social platforms from traditional, offline social contexts but also formulate an extensive program for future research. Regarding interpersonal exchanges in organizational settings, research questions of interest include how online platforms shape the exchanges between internal as well as external stakeholders such as knowledge sharing among employees (Wu, 2013), and customer engagement using online platforms (Kumar et al., 2016). Next, scholars have also attempted to extend established economic mechanisms, including matching markets (Horton, 2014; Hitsch, Hortaçsu, and Ariely, 2010) and network effects (Katona, Zubcsek, and Sarvary, 2011) to online platforms. Lastly, by their very nature, online platforms align with the study of social networks. Kane et al. (2014) have posited a series of research questions, adopting the social network analysis (SNA) lens for studying online platforms. Evidently, online platforms have opened newer research avenues across multiple fields.

The present dissertation contributes to this body of work. The increasing economic activity happening in the digital space is our primary motivation. Perhaps no other instance underscores this trend more than the online investment and financial platforms (Lee & Shin 2018). The rising prominence of these platforms is also reflected in the recent calls for research explicitly targeted towards "Fintech" (Gomber, Kauffman, & Weber 2015; Hendershott et al., 2017). Such platforms are particularly interesting because they allow users to carry out a personal activity in a highly visible and transparent setting. For instance, Venmo broadcasts a user's personal transactions, including bill and rent payments to her network. Kiva and Prosper publicize the amount of money that each investor has given to a Crowdfunding campaign (Lin & Viswanathan 2015). Social trading platforms allow users to broadcast their investment decisions in stocks and Forex (Glaser & Risius, 2017). Clearly, one observes that personal investment decisions and actions are increasingly taking place in a social domain. Given this transition, it is worthwhile to examine whether and how a focal user's investments are susceptible to other users' investment actions and decisions (Burtch, Ghose, & Wattal, 2013). The dissertation addresses this question through a series of related empirical studies.

In this introductory chapter, our objective is to situate the subsequent empirical explorations within a broader research agenda. To that end, we formulate a literature classification framework. This framework provides several advantages over other, similar efforts. Namely, our classification framework is neither constrained by a particular theoretical lens (McFarland & Ployhart, 2015) nor by the types of online platforms to which it can be applied (Kane et al., 2014). The framework comprises three categories, with each representing a different set of antecedents explaining the focal user's susceptibility to others' actions.

The rest of the chapter is structured as follows: first, we list the commonly occurring examples of information signals about others' actions, which are salient across various online platforms. In the same section, we also identify the different mechanisms that explain the link between these signals and the observer's subsequent behavior. Next, we develop the classification framework, describing some exemplary studies belonging to each of the three categories. Lastly, we map the present dissertation onto the framework, identifying the specific areas of contribution.

2. ESTABLISHING THE BASELINE

2.1 What information about others' past actions do the platforms provide?

In this section, we provide some instances of the information about others' actions as presented across different online platforms. As we note, the exact information signal is highly platform-specific. E-commerce platforms provide online reviews as well as aggregated purchase behavior of prior visitors of a product page (Chen, Wang, and Xie, 2011). On social networking platforms such as Facebook, users can see broadcast notifications as well as receive personalized messages conveying activities of their friends (Aral & Walker, 2011). On free software distribution platforms, users can see the count of prior downloads of each application (Duan, Gu, & Whinston, 2009). Along with the instances of online investment and financial platforms discussed earlier, these examples indicate that across different online platforms, contextually-relevant information about others' actions is highly salient.

2.2 What are the corresponding explanatory mechanisms?

Table 1.1 lists the different explanatory mechanisms that have appeared in the literature over the years. Each of the mechanisms explains the link between the information about others' actions and the focal user's behavior. However, the applicability of a particular mechanism is based on the study's empirical context. For instance, in pro-social contexts, information about other's prior actions creates social norms and signals socially acceptable behavior. Therefore, the observer's subsequent behavior can be explained as a function of conformity to social norms (Shang & Croson, 2009; Burtch, Ghose, & Wattal, 2013). In online retail platforms, opinions of prior customers may shape the observer's perception of the quality of the product, increasing the likelihood of future purchases (Rosario et al., 2016). Within the online platform literature, each mechanism has spawned a considerable body of work. However, it is important to note that the mechanisms are not conceptually independent in a rigid sense. Studies have frequently indicated such overlaps. For example, observational learning is rooted in the information cascade theory (Chen, Wang, & Xie, 2011; p. 240), and word of mouth is related to social contagion (Aral & Walker, 2011).

Table 1.1: Overview of the Explanatory Mechanisms

Explanatory Mechanism	**Mechanism Description**	**Examples of empirical context**
Information cascades	An individual, after observing the prior actions of others, follows their choices, *disregarding her information* as doing so is optimal in a given situation (Bikhchandani, Hirshleifer and Welch 1992)	1. Software distribution platforms (Duan, Gu, & Whinston, 2009) 2. Online retailers (Gu, Tang, & Whinston, 2013)
Conformance to norms	Information about others' prior actions signals "*the appropriate or desired behavior*," thereby guiding the observer's subsequent activity (Shang & Croson, 2009)	1. Crowdfunding platforms (Burtch, Ghose and Wattal, 2016)
Word of mouth (WOM)	It refers to the "dissemination of information (e.g., *opinions and recommendations*) through communication among people." (Chen, Wang and Xie, 2011; p. 239)	1. E-commerce platforms 2. Review platform (Rosario et al., 2016)
Observational learning	Explains the link between the "*discrete signals* expressed by the actions of other consumers *but not the reasons behind their action.*" (Chen, Wang and Xie, 2011; p. 239)	1. Music streaming platforms (Newberry, 2016) 2. Online B2B exchange platform (Koh & Fichman, 2014)

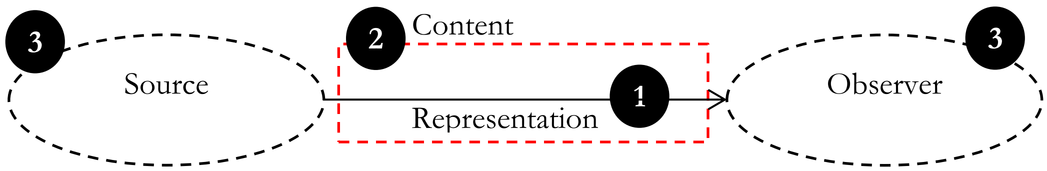
Social contagion/Peer influence	Occurs when the "actions of one's peers *change the utility one expects* to receive from engaging in a certain behavior, increasing the likelihood that one will engage in that behavior" (Aral, 2011; p. 218)	1. Social networking platforms (Aral and Walker, 2011) 2. Music streaming platforms (Bapna & Umyarov, 2015)

3. LITERATURE CLASSIFICATION FRAMEWORK

3.1 Breaking down the effect of other's actions

Having provided the examples of the information about others' actions along with the underlying mechanisms, we now proceed to develop the literature classification framework. It identifies three distinct categories of attributes, which collectively explain the susceptibility of user *i*'s (observer) behavior to user *j*'s (source) actions. First, the manner in which source's prior actions are conveyed to the observer may be pertinent. For example, information can be conveyed as generic, broadcast messages as well as through personalized directed messages (Aral and Walker, 2011). As we discuss later, the information representation may be associated with different underlying mechanisms. Second, the characteristics of the information content may explain observer's susceptibility. For instance, information that has positive connotations tends to be less influential than the information that has negative connotations (Chen, Wang, & Xie 2011). Lastly, the attributes of the source, as well as the observer, may play a role. For example, an observer may be more susceptible towards the information generated by particular sources (Forman, Ghose, & Wiesenfeld, 2008). Figure 1.1 depicts the three categories of the classification framework.

Figure 1.1: The Literature Classification Framework



1 → The behavioral impact of the different types of representation of the information

2 → The behavioral impact of the characteristics of the content of the information

3 → The behavioral impact of the characteristics of the source and observer of the information.

3.2. Comparison with other Classification Frameworks

The proposed framework offers at least three advantages over other, similar efforts. First, it applies to a wide variety of online platforms. For instance, Kane et al. (2014) limit themselves to "social media" with an emphasis on interpersonal network ties. Thus, they exclude other forms of online platforms such as "product review networks or peer production communities" (p. 280). However, our classification framework presents no such constraints because it does not assume any prior network ties between the source and the observer. Second, our framework is not wedded to any particular theoretical lens. For instance, McFarland & Ployhart (2015) draw extensively from psychology literature to distinguish social media from other forms of digital communication and derive research questions for social media research, especially in organizational contexts. However, our framework imposes no conceptual lens and hence can borrow from several theoretical disciplines. Finally, in contrast to other frameworks that emphasize exchanges between a certain kind of social actors (e.g., B2B) (Wang et al., 2017), our framework is applicable across different levels of granularity, including individuals, communities, and organizations. That being said, our framework remains limited in the research issue it addresses: how information about the source's prior actions affect the observer's subsequent behavior in online platforms.

3.3. Category 1: Attributes of the Information Representation

The first group of studies examines how different representations of the information may determine the observer's susceptibility towards a source. Notable work under this category includes the quasi-experimental study of product pages on Amazon by Chen, Wang, and Xie (2011). Authors examine two distinct ways in which Amazon conveys the information about users' behavior related to a product. Amazon provides consumer-generated reviews for each product as well as the purchase percentages based on the prior purchase actions of all users who visited the product page (p. 242). Authors argue that the two information signals distinctly affect users' subsequent purchase behavior (p. 247). The rationale for the distinct effects is that the consumer-generated reviews lead to word-of-mouth (WOM) communication while the platform-generated, aggregate information on prior purchase actions lead to observational learning (OL). In closing, authors demonstrate that different representations of information may relate to distinct explanatory mechanisms and hence, affect subsequent behavior differently.

Aral and Walker (2011) make similar arguments in their large-scale, randomized field trial. They compare the social contagion and peer effects resulting from automated broadcast notifications with those resulting from personalized referrals. The latter mode of information representation is more targeted because the source explicitly selects the users who receive such referrals (p. 1624). In automated broadcast notifications, all those in the source's network observe the information. In their study, in which both representations convey the source's adoption of a Facebook application, Aral & Walker (2011) find that broadcast notifications generate significantly stronger social contagion and peer effects, resulting in higher adoption by the observer. However, personalized referrals are more efficient, in that they lead to higher adoption per message. Thus, broadcast notifications, which require less effort from the message sender, generate more social influence but are less efficient than the personalized messages, which demand considerably more effort from the message sender. Collectively, their study shows that the effects of different information representations are distinct.

Lastly, Chen and Berger (2016) examine how the modes through which the observer acquires information may influence her subsequent sharing behavior. Specifically, an observer has a greater propensity to share "discovered" information, compared to the information she "receives from others." The underlying mechanism is that when the observer receives any information, she is more likely to evaluate its content, which depresses the propensity to share. However, when she discovered information, she is less likely to notice the underlying content variations, leading to greater sharing behavior. Consistent with this argument authors find that participants in the "Receive" condition, who were told that they had received an article through email, were more sensitive to the article's content than those who "discovered" the same article by navigating through other content. Thus, Chen & Berger (2016) show that the manner in which information reaches the observer is highly consequential. In sum, the three studies demonstrate that the information representation does affect the observer's behavioral response.

3.4. Category 2: Attributes of The Information Content

The next category refers to studies examining the attributes of the content. Typically, these studies use online reviews generated on e-commerce platforms as the empirical context (Chevalier and Mayzlin, 2006). The content of such reviews is measured in two distinct ways. Scholars measure the “quantitative surrogates” of the online reviews such as the volume and the word count (Ludwig et al. 2013; p. 87). Building on this work, studies have also explored the linguistic aspects of the review such as the level of argumentation in a review (Willemsen et at., 2011), emotions that a review conveys (Yin, Bond, & Zhang, 2014), and its affective tone (Ludwig, et al., 2013). Lastly, scholars have also examined specific product features discussed in the textual component of online reviews (Archak, Ghose, & Ipeirotis, 2011). Clearly, there exists considerable literature depicting the susceptibility of an observer’s purchase behavior to numeric and linguistic components of information about others’ product usage experiences (Floyd et al. 2014, Hong, et al. 2017).

3.5. Category 3: Attributes of The Source and The Observer

Studies categorized under this heading examine how the attributes of the source and the observer determine the observer’s susceptibility to others’ actions. One such attribute that has found repeated empirical support is the source identity. In their study of Amazon product reviews, Forman, Ghose, and Wiesenfeld (2008) showed that product reviews written by the sources who disclose their identity in the reviews are considered more helpful in making purchase decisions. A similar finding was reported by Liu & Park (2015) in their study of Yelp reviews of restaurants. In related work, Aral & Walker (2012) showed that the source’s demographic attributes such as age and gender also play an important role. Authors find that single, older males are significantly more influential as information sources on Facebook. Similarly, females have a greater influence over males than over other females. On the side of the observer as well, similar effects are found. For instance, in the same study, Aral & Walker (2012) find that younger users are systematically more affected by source's past actions. However, on the issue of gender, Aral & Walker (2012) find that women are less affected by others' actions. However, prior studies have found that others' actions have a stronger impact on women (Garbarino and Strahilevitz, 2004). In all, the demographics of the source and the observer seem to play a crucial role.

Next, research also shows that the network characteristics of the information sources, as well as their prior ties with the observer, are pertinent. For instance, Harvey, Stewart, & Ewing (2011) show that tie strength as well as the communication between the two significantly increases the propagation information, and therefore, provides greater visibility to the source's YouTube videos. These findings have found widespread empirical support. Shi and Whinston (2013) show that the effect of check-in notification from the focal user's friends on the likelihood of the user's subsequent visit to that location is higher than that of notifications from a stranger. Aral & Walker (2014) report a similar effect, suggesting that both tie strength and structural embeddedness between the source and the observer increase the behavioral impact of source's past actions.

Table 1.2 provides a summary of the framework. The present dissertation contributes to the 3rd category of studies. Specifically, we examine the collective effect of various social and economic characteristics of the source and observer on their behavior in a social trading platform. The rest of the dissertation comprises two empirical studies.

4. STRUCTURE OF THE DISSERTATION

4.1 Overview of The Chapters

In chapter 2, we model the extent to which an individual investor is susceptible to others' investment actions. Drawing from the streams of self-efficacy and audience effect, we find that a focal investor's prior mastery experiences, as well as the presence of an audience, make her less susceptible to others' actions. However, exogenous changes in uncertainty, measured using the "fear indices" (Sarwar, 2012), acts as a significant moderator as individuals discount their prior mastery experiences under higher uncertainty. However, the audience effect appears robust to the exogenous changes in uncertainty. Next, in chapter 3, we adopt a dyadic lens to explain a focal trader's preference for another specific trader. Specifically, in a directed dyad, we model trader *i*'s (an observer) preference for trader *j* (a source) for making investment decisions. We show that an observer's preference for a particular information source results from a complex interplay between economic and social cues about the source. Thus, while chapter 2 models a trader's aggregate susceptibility to others' investment actions, chapter 3 provides a more granular understanding by modeling over time variations in an observer's preference towards a specific source.

4.2 Potential Theoretical Contributions

From a theoretical standpoint, the dissertation makes several contributions, most of which are enabled by the novel features of the empirical context. First, *given the availability of longitudinal data on individual-level investment activities, we can pin down the explanatory mechanisms that the studies in other contexts mostly overlook.* For instance, in Chapter 2, we hypothesize that the presence of an audience influences the extent to which an individual's investment behavior is susceptible to others' past actions.

However, in traditional investment contexts, observing and estimating this effect is challenging as the audience of an individual investor is not readily known. Second, *we can unambiguously categorize a portion of a focal individual's activity as susceptible to others' actions.* However, in studies reviewed earlier, researchers cannot similarly dissect the behavioral outcomes, limiting their ability to model the extent to which a focal user is susceptible to others. Third, *we, at least partially, resolve several theoretical tensions that have recently emerged, especially in relation to investment behavior.* For example, in Chapter 3, we show that while the potentially favorable information about a source, such as herding (the tendency of an individual to observe and imitate the prior behavior of others) and prior payoffs, affects the observer's preference for that source, these effects operate under the constraint of physical distance (Agarwal, Catalini, & Goldfarb, 2015). In sum, the dissertation advances the current state of knowledge about human behavior, primarily related to investments.

Table 1.2: Summary of the Three Categories in the Framework

Category of studies	Central research questions	Typical Empirical Contexts	Typical measures of observer's subsequent behavior	Instances of Studies
Information Presentation	Do different modes through which a platform presents information affect its behavioral impact? Do different modes through which users acquire information affect its behavioral impact?	Facebook Amazon Controlled experiments	Adoption of a Facebook application Sales Ranks Content sharing	Aral & Walker (2011) Chen, & Wang, & Xie (2011) Chen & Berger (2016)
Information Content	Do the characteristics of the information content affect the behavioral impact of such information?	Amazon Barnes & Nobles	Sales rank of a product User-reported helpfulness of a review	Chevalier and Mayzlin (2006) Ludwig, et al. (2013) Hong, et al. (2017)
Information Source and Observer	Does heterogeneity of information sources affect the behavioral impact of information such sources generate?	Amazon Facebook MovieLens Yelp	Sales rank of a product Adoption of a Facebook application	Forman et al. (2008) Aral & Walker (2012)

Chapter 2 Doing What Others Do: Understanding Susceptibility to Ambient Social Information in Online Platforms

1. INTRODUCTION

The emergence of online platforms, which incorporate novel features for interpersonal exchanges and interactions, has created new opportunities to study human behavior (Kane et al., 2014; McFarland & Ployhart, 2015). One of the most significant developments in this regard is the platforms' ability to capture and make public social information on the actions and choices of others (Shang & Croson, 2009). For instance, on Facebook, social information is presented through notification messages that capture and broadcast personal actions of individuals to others in their networks (Aral and Walker, 2011). Such notifications often include actions, such as ordering a book on Amazon, listening to a song on Spotify, reading a news article on Slate.com or watching a movie on Netflix. Typical notifications on Facebook read, "John is watching Terminator II" and "John and Jane Doe checked in at CNNGrill at SxSW." Similarly, e-commerce platforms, such as Amazon, capture and display product browsing and purchase decisions of users to others in the form of aggregated information signals (Chen, Wang, & Xie, 2011). Social payment applications, such as Venmo and Blippy, make one's financial transactions, including online purchases, publicly visible after users link these applications with their credit cards (Rhue & Sundararajan, 2013). In sum, the emergence of online platforms has brought the once personal actions into the public domain.

We use the term ambient social information to refer collectively to social information on the actions, choices, and interactions of other users that online platforms capture and make publicly visible. We characterize the social information as being ambient to highlight the ever-present and enveloping cache of information that continually updates and supplies users with a steady stream of feeds and updates on the actions of others. Our characterization of social information as ambient is analogous to the ambiance and transparency of communication exchanges on enterprise social platforms to the employees of the firm (Leonardi, 2015).

The influence of ambient social information on the users' subsequent decisions and actions has been well established in both online as well as traditional contexts. For instance, the number of prior downloads of software tools by others increases the subsequent download activity (Duan, Gu, and Whinston, 2009). Similarly, the information about the size of the charitable contributions by prior respondents to a fundraising drive increases the subsequent donation amounts (Croson and Shang, 2013) and the magnitude of contributions by others in crowdfunding campaigns (Burtch, Ghose, Wattal, 2013) influences the behaviors of subsequent platform users. These studies highlight that the ambient social information, comprising revealed preferences of others, serves as a signal of desirability and quality.

Studies also show that the manner in which platforms present ambient social information also plays a role in determining its influence on the subsequent actions of users. For instance, Chen, Wang, & Xie (2011) demonstrate that ambient social information, when presented in the form of user-generated product reviews, affects subsequent purchase decisions differently than when it is presented as a platform-generated summary of the prior purchase decisions of other users. Recent studies suggest that users' susceptibility to ambient social information also extends to highly significant and risky decisions such as financial investments. For example, individuals are likely to buy or sell the stocks of companies based on the opinions and comments of other users on StockTwits and Twitter (Sul, Dennis, & Yuan, 2014).

Thus, the extant literature establishes that ambient social information consisting of other users' actions and decisions influences subsequent actions and decisions of the focal user. In the present study, we contribute to this literature in several ways. First, we study theoretically-motivated, user-specific idiosyncrasies that explain the extent to which a given user is susceptible to ambient social information. Thus, we address the recent calls to examine the underlying theoretical mechanisms that explain the influence of ambient social information (Bapna & Umyarov, 2015). Specifically, we find that prior mastery experiences of the focal user subsequently weakens her susceptibility to ambient social information. This effect is consistent with the prediction of self-efficacy theory (Bandura, 1977; 1997). However, we also show that the said effect is contingent on uncertainty, as a user significantly discounts her prior mastery experiences under higher uncertainty. We also find evidence for the audience effect. A focal user is less susceptible to ambient social information as her audience grows. Thus, we provide a nuanced understanding of how ambient social information and user behavior are linked in the context of online platforms.

Our study leverages a proprietary dataset obtained from a novel empirical setting of social trading platforms. These platforms typically comprise of traders, investing in stocks, and currencies. Social trading platforms continuously capture and broadcast the trading signals of the active traders. Collectively, these signals constitute the ambient social information on social trading platforms. Once a trader joins the platform, they can observe all such trading signals from other active traders, and therefore, can use such information to make investments. Most importantly, the platform provides information about the proportion of each trader's total fund that she invests by relying on others' investments actions. Therefore, by observing temporal changes in the proportion over time, we can measure the extent to which a trader is susceptible to ambient social information, namely, others' investment actions. The empirical estimation in this study is based on a dataset of investment transactions of over 12,000 traders who traded on the platform during a 46-week long observation window.

The rest of this chapter is structured as follows: in the next section, we review the background literature and derive hypotheses. Next, we provide a detailed description of the context, the dataset construction process, and the variables. The subsequent section describes the results. We also subject our findings to several robustness checks. In the penultimate section, we situate our findings in the literature on ambient social information and behavior in contexts of online platforms. We conclude the study by identifying limitations and directions for future research.

2. RELATED WORK

2.1. Ambient social information in prosocial contexts

Scholars have been interested in studying the influence of social information on subsequent behavior, especially in prosocial contexts, long before the emergence of online platforms. A prosocial context refers to a phenomenon in which individuals' actions are intended to be beneficial for others. Typical examples of prosocial contexts include charitable giving and contributions towards the public good. For instance, Reingen (1982) found that showing the list of prior individuals who complied with a request, increases the chances of the focal individual's compliance. Similarly, Frey & Meier (2004) show that University students are significantly more likely to make prosocial contributions if they are told that a majority of their peers have also contributed. Further, suggesting that social information acts as a benchmark for desirable behavior in prosocial settings, Croson & Shang (2008) show that when individuals are given information about other listeners' donations to public radio, they adjust their subsequent contributions to the levels indicated in such information. Lastly, Croson & Shang (2013) find that ambient social information has an implicit boundary condition in that if such information is too extreme, and therefore, unreliable, individuals are less affected by it.

Coming to the online platform, studies of prosocial behavior report similar results. Burtch, Ghose, & Wattal (2013) show that the aggregate contribution an online journalism project receives in a crowdfunding platform increases the subsequent contributions, indicating that individuals rely on social information about the prior lending activity (p. 511). Chen et al.(2010) and Burtch et al. (2017) show that the information about the content generated by other users increases the subsequent content generation by the focal user. In sum, in traditional as well as online prosocial contexts, ambient social information influences subsequent behavior. As the last column in Table 2.1 indicates, the underlying theoretical mechanisms that explain the influence of ambient social information include social comparison (Frey & Meier, 2004), and conformity (Croson & Shang, 2008).

2.2 Ambient social information in non prosocial contexts

Ambient social information influences subsequent behavior even in contexts that may not qualify as prosocial. The distinction between prosocial and other settings is vital because it is reflected in the underlying theories of behavior. For instance, the ideas of social contagion and peer influence operating through the revealed preferences of other users have gained considerable momentum in online contexts that are not prosocial. Bapna & Umyarov (2015) show that information about the adoption of a paid subscription to an online music service by those in an individual's network increases the chances of the subsequent adoption by that individual. Aral & Walker (2011; 2014) find that social information about an action the focal user's peers conveyed through automated notifications on Facebook, results in the greater adoption of the Facebook application by that user.

Other theoretical mechanisms that explain the influence of ambient social information on behavior in non-prosocial contexts include word-of-mouth and observational learning. These mechanisms have been tested predominantly in e-commerce settings. In such platforms, the ambient social information consists of user-generated online reviews as well as platform-generated aggregated signals about prior purchases (Archak, Ghose, & Ipeirotis, 2011; Chen, Wang, & Xie, 2011). While the baseline effect of ambient social information influencing subsequent behavior is persistent (Chevalier & Mayzlin, 2006), research also shows that the manner in which the ambient social information is presented to the users is highly consequential. Chen, Wang, & Xie (2011) show that the user-generated reviews, which reflects the word-of-mouth mechanism, drive subsequent behavior very differently than platform-generated information on aggregate activities of other users, which reflects observational learning.

In conclusion, the extant literature establishes that ambient social information influences users' subsequent actions and decisions. In Table 2.1, we categorize the studies based on their treatment of user-level heterogeneity. While a handful of studies explicitly model the heterogeneity, the majority of the work abstracts it away, providing only a macro-understanding of susceptibility to ambient social information (Table 2.1). However, it also presents several avenues for future research not the least of which is the lack of examination of user-specific idiosyncrasies which may affect the extent to which a user is susceptible to ambient social information. Most studies mentioned so far mask the heterogeneity across users, thereby assuming that all users are uniformly susceptible to ambient social information.

This assumption, however, may not always be valid as indicated by a handful of studies that examine user-specific idiosyncrasies. Mostly focusing on product adoption as the outcome behavior, these studies show that a person's susceptibility to ambient social information is contingent on their demographic attributes such as age and gender (Aral & Walker, 2012; Goel & Goldstein, 2014), users' network characteristics such as tie strength and structural embeddedness (Aral & Walker, 2014), and identity mechanisms such as identity congruence and identity esteem (Shang, Reed, & Croson, 2008). Thus, the present understanding of the behavioral implications of ambient social information can be further enriched by incorporating user-specific idiosyncrasies and heterogeneity. This issue is particularly true in contexts that are not prosocial and therefore are less driven by the users' conformity to social norms. In the present study, we address this gap by examining two user-specific attributes, rooted in established social psychological theories, namely self-efficacy (Bandura, 1977; 1997) and audience effect (Bond, 1982).

Table 2.1: Classification of Relevant Literature

	User-level heterogeneity modeled	User-level heterogeneity abstracted
Prosocial context	Shang, Reed, & Croson (2008) Reingen (1982)	Frey & Meier (2004), Croson & Shang (2008) Chen et al. (2010) Shang & Croson (2013) Burtch, Ghose, & Wattal (2014) Burtch et al. (2017)
Non-prosocial context	Aral & Walker (2012; 2014) Goel & Goldstein (2014)	Chevalier & Mayzlin (2006) Duan, Gu, & Whinston (2009) Aral & Walker (2011) Chen, Wang, & Xie (2011) Bapna & Umyarov (2015)

3. STUDY CONTEXT

In this section, we discuss the empirical context of the study to clarify the link between the above discussion and the subsequent formulation and testing of hypotheses. We use a social trading platform as the empirical context. In general, financial investment platforms are increasingly adopting features of online social platforms to improve an individual's decision-making (Wohlgemuth, Berger, & Wenzel, 2016). These "Fintech" platforms operate in a myriad of ways, ranging from the provision of benchmarks for investment opportunities (e.g., OpenFolio) to allowing users to expose their portfolio to other trader's investment decisions (e.g., ZuluTrade). Social trading platforms are of the latter type because they allow traders to set up automatic replication of other traders' investment signals (Doering, Neumann, & Paul, 2015). We refer to this platform as XTrader (a pseudonym). The platform enables users around the world to invest in leading company stocks, foreign currencies, and commodities.

In order to join the platform, a user is required to have a certain minimum balance to start the trading activity. Upon joining, the platform allows users to trade through a variety of instruments. The platform continuously captures and broadcasts the trading activities of active traders, generating a digital stream of social information. This information is readily observable on each trader's profile as well as through results of a configurable tool for trader-search (see Appendix I). Through either channel, a trader can observe the ambient social information, consisting of other traders' investment signals, and choose to rely on it (or not) in making their investments.

Social trading platforms, including the one we study, offer a novel mechanism through which users can leverage the ambient social information, including the other traders' investment signals. This mechanism, termed as copy trading, operates as follows: a trader *i* can observe the trading signals of trader *j*. Then, *i* can allocate a portion of her fund to *j*. After the allocation, all the currently open and subsequent investments by *j* get automatically replicated on trader *i* 's portfolio. Thus, by observing a trader's aggregate allocation to other traders, one can model the extent to which trader *i* is susceptible to ambient social information.

4. HYPOTHESIS DEVELOPMENT

4.1. Self-efficacy belief and susceptibility to ambient social information

Having reviewed the relevant literature and the empirical setting of the study, we next formulate a series of hypotheses, predicting a user's susceptibility to ambient social information. In the first hypothesis, we draw on the self-efficacy literature as the explanatory theoretical mechanism. Since Bandura's (1977; 1997) seminal work, self-efficacy has received considerable research attention. While the concept has been described in several ways (e.g., creative self-efficacy by Tierney & Farmer, 2002), its definition has remained mostly unchanged. It is defined as the person's belief that they “possess the competence necessary to be effective and influential in one’s environment” (Fast, Burris, & Bartel, 2014; p. 1014). Studies have shown that when an individual perceives herself as efficacious, her behavior changes significantly (Liu et al., 2014; Hartog & Belschak, 2012).

We propose that a user’s self-efficacy beliefs determine her susceptibility to ambient social information. A person with low self-efficacy belief will “seek out relationships with those who advance their goals and avoid those who obstruct their goals” (Shea, Davisson, & Fitzsimons, 2013; p. 1032), indicating a potentially compensatory behavior. That is, a user may look to compensate for their lack skills and expertise by relying on others’ assistance and help. Applying this argument to ambient social information, we claim that a user with low self-efficacy will exhibit greater compensatory behavior and therefore, higher susceptibility to ambient social information than a user with high self-efficacy.

There are several sources of self-efficacy belief, most effective of which is “the mastery experiences or repeated performance accomplishments” (Boyd & Vozikis, 1994; p. 67). Such experiences are most effective in enhancing self-efficacy beliefs because “they provide the most authentic evidence of whether one can muster whatever it takes to succeed” (Bandura, 1997; p. 80). By this argument, favorable outcomes of prior actions may act as mastery experiences for an individual, thereby strengthening her self-efficacy belief and weakening her susceptibility to ambient social information. Therefore, our first hypothesis is as follows:

<u>Hypothesis 1 (H1):</u> Mastery experiences from prior actions by a user will be negatively related to her susceptibility to ambient social information.

4.2 Moderating influence of uncertainty

In the second hypothesis, we examine the moderating role of uncertainty. Borrowing from Walker et al. (2003), we consider uncertainty as "the variability inherent to the system under consideration" (p. 4), namely in the financial markets involving company stocks and currencies. Numerous studies in behavioral decision theory suggest that uncertainty significantly alters individuals' decision making (Kahneman & Tversky, 1979; Halter & Dean, 1971). Specifically, as uncertainty increases and tasks become difficult, individuals experience self-doubt (Oleson et al. 2000) as well as increasingly rely on others' advice (Gino & Moore, 2007). Similar arguments have been made in studies of social information in traditional, offline settings. For example, Shang & Croson (2009) argue that "individuals are more likely to be positively influenced by social information when the situation is (seen as) ambiguous" (p. 1426). In sum, uncertainty represents an offsetting mechanism to that predicted in H1. While the mastery experiences from prior actions are likely to weaken a user's susceptibility to ambient social information, increase in uncertainty is expected to push a user towards relying more on ambient social information.

Extant literature provides some indication that under uncertainty, sources of self-efficacy have weaker impacts. For example, in group settings, Gibson (1999) shows that under high uncertainty, a "group is not sure how it achieved prior outcomes" (p. 140). As a result, uncertainty weakens the effect of the group's self-efficacy. Bandura (1997) also argues that the consequences of self-efficacy sources, specifically the importance of mastery experiences, are contingent on the exogenous, environmental conditions. Therefore, we argue that under higher uncertainty, a user with mastery experiences, and therefore higher self-efficacy belief, will be more susceptible to ambient social information.

Uncertainty may moderate the effect of prior mastery experiences through another route. Specifically, increasing uncertainty may create a “level playing field,” as individuals across the skill-levels may become unsure of the paths to achieve a favorable outcome (Boudreau, Lacetera, & Lakhani, 2011). In other words, increase in the uncertainty may reduce the overall activity levels across the platform. In our context, this effect will translate into a reduction in aggregate allocation because, under higher uncertainty, traders may wish to hold back their investment funds. This overarching effect may render prior mastery experiences redundant, leading to a substitution effect. In either way, we expect that as uncertainty increases, the effect hypothesized in H1 will weaken. Thus, our second hypothesis is as follows:

Hypothesis 2 (H2): Increase in uncertainty will weaken the negative relationship between the user’s prior mastery experiences and her susceptibility to ambient social information.

4.3 Audience effect and susceptibility to ambient social information

Our third hypothesis postulates the relationship between the presence of an audience and the user’s susceptibility to ambient social information (Pfattheicher & Keller, 2015). Recent research shows that an audience significantly influences a user’s behavior in online settings. Bapna et al. (2016) show that in online dating platforms, the veil of anonymity makes individuals more uninhibited as indicated by the diversity (e.g., interracial mates, same-sex mates) of the profiles they visit. However, when users cannot visit others’ profiles anonymously, their visits become more conventional. Huang, Hong, & Burtch (2016) find that the characteristics of a user’s online reviews on Yelp and

TripAdvisor become increasingly positive when such reviews attain a larger audience through the integration of these platforms with Facebook. Collectively, these effects show that user behavior becomes increasingly conforming when the users are under observation.

In online platforms, the presence of an audience indicates a dilemma for the focal user. On the one hand, platforms provide an opportunity for the individual to observe and benefit from others' actions through ambient social information. However, because a user's decision to rely on others is itself transparent, it is not without a social cost. The underlying reasoning for such costs comes from the extant social psychological literature on self-presentation, which argues that in the presence of an audience, individuals want to project themselves in a favorable light. There are at least two alternative explanations for this effect. Baumeister (1982) states that self-presentation stems from the person's expectation of rewards from a "pleased" audience as well as the urge to present a public self that is consistent with one's ideal self (p. 3). On the other hand, Bond (1982) claims that presence of audience creates a fear of embarrassment and loss of face, causing people to engage in self-presentation. However, regardless of the reasoning, "the performer will be motivated to project an image of competence in the presence of others" (Bond & Titus, 1983; p. 267).

By drawing from these arguments, we claim that given the highly transparent nature of online platforms, a user's susceptibility to ambient social information may be viewed unfavorably by her audience (Lee, 2002). As a result, the bigger the audience a user has, the more socially costly it will be for her to rely on ambient social information. That is, as the audience grows, the self-presentation mechanism will overshadow the benefits of relying on the ambient social information. Thus, our final hypothesis is as follows:

Hypothesis 3 (H3): A user's audience will be negatively related to her susceptibility to ambient social information.

5. DATASET CONSTRUCTION AND VARIABLE DESCRIPTION

5.1 Dataset construction

Our original dataset comprises of 15,273 traders observed over 46 consecutive weeks. By pairing each trader with each observation week (i.e., balanced panels), we obtain 702,558 observations. From this dataset, we exclude approximately 10% observations in which the aggregate allocation of the trader towards others exceeded the 100% threshold because such observations may constitute measurement error. Retaining these observations produces no meaningful impact on the sign and the magnitude of the eventual estimates. After excluding such records, the observation count reduces to 683,357. The count of traders remains the same (15,273). However, using the entire dataset for estimation may not be valid. Specifically, at any point in time, some traders may be dormant for several, unobserved circumstances. To address this concern, we only include observation for trader *i* in week *t* if at least one transaction was posted to *i*'s account in week *t*. Thus, we consider the presence of transaction on a trader's account as the uniform baseline. After applying this filter, the dataset reduces to 71,358 observations, consisting of 12,605 traders, which we use for hypotheses testing. In this dataset, each panel comprises of a different trader observed over time.

5.2 Variable descriptions

Dependent Variable: The primary objective of the study is to model the extent to which individuals are susceptible to ambient social information. Thus, as the outcome variable, for each trader *i* in week *t*, we calculate the total percentage of the fund she allocated to others (SUM_ALLOCATION). For instance, suppose traders *i* and *j* have respectively allocated *x%* and *y%* of their funds to others in week *t*. If $x > y$, then we claim that trader *i* is more susceptible to ambient social information than trader *j* because a greater portion of trader *j*'s portfolio is under her control.

Independent Variables: Our first independent variable captures a trader's prior mastery experiences which in turn are expected to bolster her self-efficacy beliefs (Bandura, 1997). In social trading, prior positive investment gains represent such experiences. Thus, for each trader *i* observed in week *t*, we calculate the count of weeks in which *i* made profits through independent transactions, starting from the first week in which a transaction was reported on *i*'s portfolio till week *t* (PROFITABLE_WEEK_COUNT). For instance, assume that during the 46-week long observation window, the first transaction was posted to trader *i*'s portfolio in week t_i. Thus, the value of PROFITABLE_WEEK_COUNT for trader *i* in a subsequent week $t_i + k$ is the count of intermediate weeks in which *i* made a profit through independent transactions. Thus, in any week, this variable captures the aggregate instances of favorable, prior mastery experiences that trader *i* has attained.

Our second independent variable is related to the audience effect (H3). We compute this measure as follows: for each trader *i* in week *t*, we count the distinct, other traders who have sent messages to *i* between the first week from the observation window on which a transaction was posted to *i*'s portfolio and week *t* (MESSAGE_SENDERS).

Thus, the computation approach for this variable is consistent with that for the PROFITABLE_WEEK_COUNT. We argue that the higher the count of message senders, the larger is the audience of the focal trader. To suppress endogeneity between the independent variables and the trader's susceptibility in the present week, we incorporate a 1-week lag for the former. Also, we log-transform the two independent count variables after incrementing their original values by 1.

Moderating Variable: For measuring uncertainty (H2), we use the Chicago Board of Exchange's Volatility Index (VIX™). The index predicts the fluctuations in S&P 500 stocks. While the index is based on large companies listed on American stock markets, it is known to have global ramifications (Sarwar, 2012). Over the years, the accuracy of VIX™ based forecasts of market volatility has increased (Corrado & Miller, 2005). We compute the standard deviation VIX™ values for each observation week (SD_VIX). We argue that higher the standard deviations in VIX™ higher the uncertainty. The choice of measure is consistent with the existing literature on uncertainty in financial investments (Chan-Lau, Liu, & Schmittmann, 2015).

Control Variables: Trading strategies could be driven by several demographic factors such as trader's nationality (Beugelsdijk & Frijns, 2010) and gender (Sundén & Surette, 1998). To account for such unobserved, time-invariant sources of unobserved heterogeneity, we incorporate trader fixed effects during the estimation. Next, a focal trader's investments may also be a function of unobserved shocks to the system during the observation window. To account for the effect of such shocks on trading behavior, we include 45 weekly dummies. Also, we incorporate two time-varying attributes. First, we control for the trader's tenure on the platform. We compute tenure as the count of days between the trader's platform joining date and the first date of a given observation week. We subject this count variable to a log transformation after incrementing it by 1 (LOG_TENURE). Note that a trader's joining date may fall outside the 46-week observation window. Using tenure as a control variable, we account for any unobserved heterogeneity in investments that may result from a given trader's exposure to the platform. Next, we also control for the trader's gain through copy-trading in each week. Specifically, we sum gains that trader *i* made in a week *t* (COPIED_GAIN). We reason that a trader who has profited by relying on another trader's investment signals may be systematically more susceptible to ambient social information in the present week. For the estimation, we lag the COPIED_GAIN control variable by 1 week. Table 2.2 presents the pairwise correlations and Table 2.3 provides descriptive statistics. We observe that the VIF values are proximal to acceptable thresholds (Rai et al., 2015). Also, all the pairwise correlation coefficients are below 0.5. Therefore, we conclude that the variables do not suffer from any major concern pertaining to collinearity.

Table 2.2: Pairwise Correlations and Collinearity Diagnostics

	V1	V2	V3	V4	V5	V6
Sum of Allocation (V1)	1					
Log of Tenure (V2)	-0.0301*	1				
Total Gain from Copying (V3)	0.0545*	-0.0049	1			
Log of Profitable Week Count (V4)	-0.1315*	0.1111*	-0.0019	1		
Log of Message Senders (V5)	-0.0756*	0.1504*	0.0097*	0.3979*	1	
Std. Dev of VIX (V6)	-0.2727*	0.0464*	-0.0463*	0.0321*	-0.0184*	1

* $p < 0.05$
Highest VIF Score = 1.21
Mean VIF Score = 1.11

Table 2.3: Descriptive Statistics

Statistics	Sum of Allocation	Log of Tenure	Total Gain from Copying	Log of Profitable Week Count	Log of Message Senders	Std. Dev. of VIX
Average	71.08209	6.133425	0.126421	0.254136	0.325687	0.630272
Minimum	0	0	-89.1643	0	0	0.08204
Maximum	100	7.724005	33.21812	2.772589	8.540323	2.44788
Std.Deviation	43.69851	0.595932	2.154948	0.484943	0.820481	0.538581

6. ANALYSIS

6.1 Hypothesis testing

We begin hypothesis testing with the sum of allocation by a trader in a week as the outcome measure. To account for any unobserved heterogeneity across the traders, we incorporate trader fixed effects. Moreover, in each model, we include a vector of fixed effect dummies for the observation weeks. We incrementally estimate the coefficients of interest, starting with a baseline model (Model 1) which computes the estimates for the time-varying two control variables: log of trader's tenure and the average gains per copied transaction (Table 2.4). As expected, we find that an increase in the gains through copying increases total allocation in the subsequent weeks. Next, in Model 2 and Model 3, we respectively incorporate the measures for testing H1 (count of profitable weeks) and H3 (count of distinct senders of messages). In Model 4, we simultaneously estimate the two main effects. Next, for testing H2, we incorporate the interactions between uncertainty and the count of profitable weeks (Model 5). Subsequently, we introduce the interaction between uncertainty and the audience effect (Model 6). Although we have not hypothesized this interaction effect, it is worthwhile assessing whether the "watching eyes" phenomenon is contingent on the uncertainty. Lastly, we test all the hypothesized effects simultaneously in Model 7. In this fully specified model, we find that the three hypothesized effects are jointly significant.

Coming to the effect sizes, we find (Model 4) that a 1% increase in the Trader's mastery experiences reduces the total allocation in the subsequent week by 2.57 (H1). A similar increase in the count of message senders reduces the subsequent allocation by a factor of 1.46 (H3). Next, we find that the interaction between the uncertainty and prior mastery experiences operates expectedly. To further clarify the nature of the interaction term hypothesized in H2, we plot the marginal effect of the count of profitable weeks for varying levels of uncertainty (Figure 2.1). The figure shows that under conditions of high uncertainty *(red line)*, prior mastery experiences have a marginally negative effect on total allocation. However, under conditions of low uncertainty *(blue line)*, the negative effect of prior mastery experiences on total allocation is stronger. This finding is consistent with H2.

Figure 2.1: Interaction between Mastery Experiences and Uncertainty

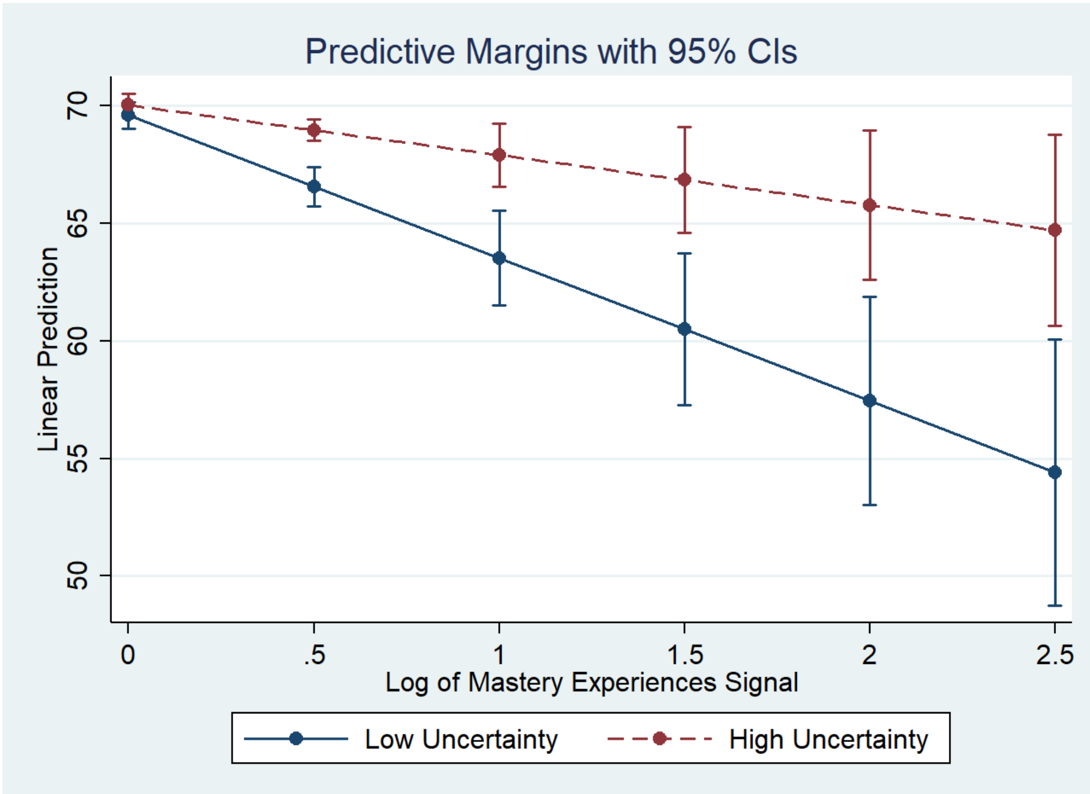


Table 2.4: OLS with Trader and Time Fixed Effects

(DV = Sum of Allocation)

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**
Trader Experience (Log)	4.148^{+}	4.027^{+}	4.113^{+}	4.028^{+}	3.987^{+}	4.048^{+}	4.005^{+}
	(2.358)	(2.348)	(2.348)	(2.342)	(2.339)	(2.339)	(2.338)
Sum of Copied Gain	0.161**	0.147**	0.149**	0.141**	0.144**	0.144**	0.145**
	(0.0516)	(0.0515)	(0.0517)	(0.0516)	(0.0517)	(0.0517)	(0.0517)
Count of Profitable Weeks (Log)		-3.392***		-2.570**	-4.339***	-2.608**	-4.097***
		(0.923)		(0.932)	(1.011)	(0.933)	(1.018)
Count of Message Senders (Log)			-1.781***	-1.462**	-1.407**	-2.162***	-1.811**
			(0.499)	(0.508)	(0.507)	(0.555)	(0.563)
Count of Profitable Weeks (Log)×SD_VIX					2.318***		1.972***
					(0.382)		(0.408)
Count of Message Senders (Log)×SD_VIX						1.173***	0.663*
						(0.285)	(0.302)
Constant	64.37***	65.40***	64.59***	65.33***	65.70***	65.24***	65.59***
	(13.98)	(13.92)	(13.92)	(13.88)	(13.87)	(13.87)	(13.86)
N	58753	58753	58753	58753	58753	58753	58753
Adj. R^2	0.7248	0.7250	0.7251	0.7252	0.7255	0.7253	0.7255
AIC	517731.0	517683.7	517671.4	517647.0	517588.6	517617.8	517582.0
BIC	518144.1	518105.8	518093.5	518078.1	518028.7	518057.9	518031.0
Trader Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y	Y	Y	Y	Y

Robust standard errors clustered on traders are reported

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.2 Robustness Check-I: Are the results robust to the choice of measures for the variables of interest?

The results obtained so far indicate empirical support for the hypothesized effects. Moreover, the results are robust to the outcome measure as well as the estimation approach. In this section, we subject our findings to additional robustness checks. We begin by assessing whether the results are robust to the choice of the uncertainty measures (Table 2.5). We replace the standard deviation of VIX with the standard deviation of VXO, which is an alternate index by Chicago Board of Exchange. Much like VIX, VXO is also conceived as a fear index, but while VIX is calculated using the S&P 500 Index, VXO is calculated using S&P 100 Index (Guo & Wohar, 2006). In addition, we also use the standard deviation of VXST, which represents a short-term uncertainty in the financial markets. We find that after changing the measure of uncertainty, the key effects are retained (Model 8-9). Therefore, we conclude that the results are robust to the choice of the measures.

Table 2.5: Re-estimation using Alternate Measures

(DV = Sum of Allocation)

	Model 8	**Model 9**
Trader Experience (Log)	4.019^{+}	4.000^{+}
	(2.339)	(2.338)
Sum of Copied Gain	0.145^{**}	0.145^{**}
	(0.0517)	(0.0518)
Count of Profitable Weeks (Log)	-3.995^{***}	-3.879^{***}
	(1.010)	(1.001)
Count of Message Senders (Log)	-1.686^{**}	-1.767^{**}
	(0.558)	(0.557)
Count of Profitable Weeks (Log)×SD_VXST	1.119^{***}	
	(0.229)	
Count of Message Senders (Log)×SD_VXST	0.281^{+}	
	(0.170)	
Count of Profitable Weeks (Log)×SD_VXO		1.798^{***}
		(0.384)
Count of Message Senders (Log)×SD_VXO		0.620^{*}
		(0.292)
Constant	65.47^{***}	65.59^{***}
	(13.86)	(13.86)
N	58753	58753
Adj. R^2	0.7255	0.7255
AIC	517592.2	517586.3
BIC	518041.3	518035.4
Trader Fixed Effects	Y	Y
Observation Week Fixed Effect	Y	Y

Robust standard errors clustered on Traders are reported
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.3 Robustness Check-II: Are the results in the present week driven by the high performing traders from the prior week?

A possible concern is that traders who were the highest gain in week *t-1* may have no better-performing traders to allocate funds in week *t*. As a result, the observed reduction in SUM_ALLOCATION may be driven by the presence of such high-performing traders. To test this possibility, we first sort all traders in a week based on their total, independent gain. Consider trader *i* who achieved gain G_i in week *t-1*. If no other trader produced gain greater than G_i in *t-1* then, after sorting, trader *i* appears in the first position in week *t-1*. Next, we exclude top 10, 25 and 50 traders from each week. Thus, we can reasonably argue that the remaining traders in each week had an option of copying the investment signals of a better-performing trader. Consequently, the outcome variable values for the retained traders can be construed as a deliberate choice. After removing the high performing traders from each week, we find consistent results (Model 10-12). Hence, we conclude that the results are not contingent on high-performing traders from prior weeks (Table 2.6).

Table 2.6: Excluding High Performing Traders from Prior Weeks

(DV = Sum of Allocation)

	Model 10	**Model 11**	**Model 12**
Trader Experience (Log)	4.536^{+}	4.891^{+}	5.087^{*}
	(2.446)	(2.531)	(2.536)
Sum of Copied Gain	0.171^{***}	0.164^{**}	0.156^{**}
	(0.0509)	(0.0516)	(0.0518)
Count of Profitable Weeks (Log)	-5.715^{***}	-5.454^{***}	-5.248^{***}
	(1.032)	(1.065)	(1.080)
Count of Message Senders (Log)	-2.071^{***}	-1.916^{**}	-2.018^{**}
	(0.582)	(0.608)	(0.622)
Count of Profitable Weeks (Log)×SD_VIX	2.551^{***}	1.970^{***}	1.796^{**}
	(0.437)	(0.531)	(0.550)
Count of Message Senders (Log)×SD_VIX	0.734^{*}	0.494	0.512
	(0.366)	(0.412)	(0.429)
Constant	62.40^{***}	60.06^{***}	58.86^{***}
	(14.47)	(14.96)	(14.99)
N	51791	47668	46884
Adj. R^2	0.7598	0.7422	0.7468
AIC	451140.5	415522.1	407813.7
BIC	451459.2	415785.2	408032.6
Trader Fixed Effects	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y

Robust standard errors clustered on Traders are reported
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.4 Robustness Check-III: Are the results contingent on the structure of the dataset

Next, we test whether the statistically significant results are consistent across different, randomly generated subsamples of the original dataset (Table 8). Estimating effects with random subsamples assesses whether the significance detected is mostly a function of the sample size (Yao et al., 2009). More specifically, a typical concern about large-sample, observational studies is the possible detection of small differences in means, which may not be practically meaningful (Lin, Lucas, & Shmueli, 2013). To address this concern, we randomly prune our dataset by 10%, 20%, 30%, 40%, and 50%. We rerun the OLS estimator with trader and observation week fixed effects. In Models 13-17, we find that across the five subsamples results remain consistent (Table 2.7). Thus, we conclude that the support for the hypothesized effects is not contingent on the configuration of the original dataset.

Table 2.7: Re-estimation using Randomly Generated Subsamples

(DV = Sum of Allocation)

	Model 13	**Model 14**	**Model 15**	**Model 16**	**Model 17**
Trader Experience (Log)	3.848	4.899*	4.572^{+}	1.956	5.129^{+}
	(2.418)	(2.493)	(2.617)	(2.732)	(3.016)
Sum of Copied Gain	0.115*	0.120^{+}	0.159*	0.115^{+}	0.107
	(0.0542)	(0.0612)	(0.0625)	(0.0642)	(0.0680)
Count of Profitable Weeks (Log)	-3.902***	-4.629***	-4.107***	-2.537*	-3.420**
	(1.025)	(1.138)	(1.186)	(1.236)	(1.325)
Count of Message Senders (Log)	-1.764**	-2.117***	-2.794***	-2.008**	-1.402*
	(0.579)	(0.616)	(0.662)	(0.697)	(0.683)
Count of Profitable Weeks (Log)×SD_VIX	1.795***	2.003***	1.638***	1.879***	1.753**
	(0.427)	(0.470)	(0.473)	(0.535)	(0.546)
Count of Message Senders (Log)×SD_VIX	0.824**	0.800*	1.006**	0.972*	0.200
	(0.317)	(0.352)	(0.379)	(0.383)	(0.392)
Constant	66.20***	60.16***	62.26***	77.45***	57.84**
	(14.34)	(14.78)	(15.51)	(16.19)	(17.89)
N	52940	47012	41044	35313	29328
Adj. R^2	0.7293	0.7308	0.7331	0.7392	0.7439
AIC	464687.5	411459.6	357562.8	305698.4	251837.3
BIC	465131.4	411897.5	357993.9	306122.0	252251.6
Trader Fixed Effects	Y	Y	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y	Y	Y

Robust standard errors clustered on Traders are reported

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.5. Robustness Check IV: Fractional Response Model

Next, we estimate whether the results change if we replace the OLS estimator with one designed specifically for fractional response outcomes (Papke & Wooldridge, 2008). These estimators are useful as they enforce the [0,1] bounds on the post-estimation predictions. Although using a fractional response model for unbalanced panels is a topic of future research for econometricians (Papke & Wooldridge, 2008), results obtained with fractional estimator without Trader-level fixed effects may still indicate the robustness of the findings to alternate specifications. Accordingly, we use the fractional response estimator with dummies for observation week (https://www.stata.com/manuals/rfracreg.pdf) (Model 18-20). In Model 18, we compute heteroskedastic-consistent standard errors, in Model 19, we compute Bootstrap standard errors with 100 repetitions, and in Model 20, we cluster the standard errors on Traders. We consistently obtain similar estimates for the three hypothesized effects, suggesting that the findings are stable to the choice of the estimator (Table 2.8).

Table 2.8: Fractional Response Model

(DV = Sum of Allocation/100)

	Model 18	**Model 19**	**Model 20**
Trader Experience (Log)	-0.0723**	-0.0723**	-0.0723*
	(0.0240)	(0.0225)	(0.0345)
Sum of Copied Gain	0.0310***	0.0310***	0.0310***
	(0.00459)	(0.00482)	(0.00470)
Count of Profitable Weeks (Log)	-0.970***	-0.970***	-0.970***
	(0.0398)	(0.0426)	(0.0528)
Count of Message Senders (Log)	-0.224***	-0.224***	-0.224***
	(0.0226)	(0.0224)	(0.0294)
Count of Profitable Weeks (Log)×SD_VIX	0.256***	0.256***	0.256***
	(0.0532)	(0.0473)	(0.0490)
Count of Message Senders (Log)×SD_VIX	-0.00885	-0.00885	-0.00885
	(0.0317)	(0.0328)	(0.0293)
Constant	2.222***	2.222***	2.222***
	(0.148)	(0.141)	(0.209)
N	58753	58753	58753
AIC	38620.8	38620.8	38620.8
BIC	39078.8	39078.8	39078.8
Trader Fixed Effects	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y

Model 18: Heteroskedastic consistent standard errors
Model 19: Bootstrap standard errors with 100 repetitions
Model 20: Standard errors clustered on Traders
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.6 Robustness Check V: Panels with varying outcome

Finally, we test whether the results are driven by the panels in which the outcome variable does not vary throughout the observation window. Accordingly, we compute the standard deviation of SUM_ALLOCATION in each panel. We exclude panels in which the standard deviation is zero. In the dataset, we find 5,013 such panels (N = 13,640). After excluding these observations, we obtain consistent support for the hypothesized effects (Model 21, Table 2.9). Thus, we conclude that the findings are robust to the panels in which the outcome does not vary over the observation window.

Table 2.9: Excluding Panels in which Outcome Variable does not Vary

(DV = Sum of Allocation)

	Model 21
Trader Experience (Log)	4.859^{+}
	(2.596)
Sum of Copied Gain	0.136^{*}
	(0.0552)
Count of Profitable Weeks (Log)	-4.663^{***}
	(1.154)
Count of Message Senders (Log)	-1.944^{**}
	(0.656)
Count of Profitable Weeks (Log)×SD_VIX	1.073^{**}
	(0.414)
Count of Message Senders (Log)×SD_VIX	0.309
	(0.326)
Constant	62.96^{***}
	(15.48)
N	50216
Adj. R^2	0.7362
AIC	448283.1
BIC	448724.3
Trader Fixed Effects	Y
Observation Week Fixed Effects	Y

Robust standard errors clustered on Traders are reported
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7. CONCLUSION

We began the study by observing that while the extant literature establishes the link between ambient social information and subsequent behavior in online platforms, it leaves several questions unanswered, namely the lack of examination of user-specific characteristics and heterogeneity. This issue is even more pertinent when the context is not prosocial, and hence, does not enforce conformity to social norms. Our study addresses this gap by hypothesizing and demonstrating effects that are consistent with social psychological theories. Thus, our study opens the black box of ambient social information and behavior, thereby expanding on the known theoretical mechanisms available to scholars of user behavior in online platforms. Future studies can build on our work by examining other, theoretically-motivated attributes that vary across users and therefore, are likely to affect users' susceptibility to ambient social information.

Our study also provides a comparative account of different categories of theoretical mechanisms. Self-efficacy and the audience effect have developed as distinct research streams in psychology literature. While self-efficacy is based on an individual's belief in their competence (Bandura, 1977, 1997) and is inherently personal, self-presentation operates only in the presence of other individuals (Bond, 1982) and therefore is decidedly social. Consequently, self-efficacy and audience effects can be viewed as instances of two distinct classes of antecedents of human behavior: personal and social. Thus, in the context of online platforms, our study compares the effect of the two kinds of mechanisms on individual behavior.

We also contribute to the self-efficacy literature. The variety of outcomes that self-efficacy influences range from making creative contributions (Richter et al., 2012) to acting as a whistleblower (MacNab & Worthley, 2007). In the information systems (IS) literature, the notion of self-efficacy has been predominantly used in assessing technology usage (Thatcher & Perrewé, 2002). In online contexts, studies have shown that self-efficacy positively influences contributions to a public knowledge repository (Ray, Kim, & Morris, 2014). We advance this literature by exploring the link between sources of self-efficacy beliefs and a user's susceptibility to ambient social information. From a methodological standpoint, our study is consistent with recent calls to conduct within-person studies to gain a better understanding of how self-efficacy operates (Yeo & Neal, 2013). We use the count of prior profitable weeks based on self-trades as the measure for prior mastery experiences, which we suggest enhance self-efficacy beliefs. It may be that these self-trades by the individual were based on observations of other traders rather than based on their own judgment. If this is the case, the true level of self-efficacy is potentially lower than what our measure suggests. Yet, the data suggests that count of prior profitable weeks provides theoretically consistent results, lending credence to our approach of using this signal as an indirect measure for assessing this social-psychological construct.

Our study is an important step forward towards understanding audience effect. Given that the online platforms increase the transparency of user behavior, the audience effect is especially pertinent in online platforms. While it is known that in the presence of audience, a focal user follows an expected pattern of behavior such as visiting online dating profiles (Bapna et al., 2016) and contributing online reviews (Huang, Hong, & Burtch, 2016), we show that the audience effect also influences a user's susceptibility to ambient social information and nudges her into making autonomous decisions. More generally, we find evidence that being watched makes a user less susceptible to others' actions and decisions. Future work can examine the generalizability of this finding to other online platforms.

Lastly, our study contributes to the understanding of human behavior in the nascent phenomenon of social trading. In the last decade, technologies have begun transforming even the most fundamental aspects of financial and investment markets (Morris, 2016). While there have been some attempts at understanding how individuals make investment decisions on these platforms (Wohlgemuth, Berger, & Wenzel, 2016), there is no systematic inquiry about the underlying social psychological mechanisms that explain trader behavior. Our study addresses this gap by obtaining theoretically consistent results.

Predicting a trader's susceptibility to ambient social information has considerable practical implications for the platform's business. Typically, a social trading platform relies heavily on the presence of traders who make profitable trades and therefore, can attract new users as well as generate money for the platform (Doering, Neumann, & Paul, 2015). Thus, most social trading platforms, including XTrader, seek to retain such traders. Currently, however, most social trading platforms have static, pre-determined incentive schemes that reward traders as and when they clear different performance thresholds. However, as we have shown, under higher uncertainty, even the traders who have made profitable self-trades, become more susceptible to others' trading signals as well as reducing self-trades. Given this insight, social trading platforms can devise incentive schemes that may provide an additional reward to the profitable traders for the self-trades made under higher uncertainty, thereby motivating such traders not to discount their expertise.

Our study is not without limitations. Most importantly, in order to establish the validity of the measures and the hypothesized mechanisms, we rely on the platform's novel feature of unambiguously attributing each transaction either to the focal trader's autonomous decision or her susceptibility to others' trading signals. However, we have no direct measures of user-specific attributes of self-efficacy and audience effect as in most other studies? (Thatcher & Perrewé, 2002). Consequently, while our study finds theoretically consistent results, we cannot make more definitive claims. For instance, self-efficacy can be classified into general and specific efficacy (Agarwal, Sambamurthy, & Stair, 2000) and these more granular types may have a different influence on dependency behavior. Future work can measure the specific aspects self-efficacy in online contexts to develop and test more granular hypotheses.

In conclusion, our study shows that although online platforms generate ongoing information about actions and revealed preferences of the users, social psychological mechanisms play a crucial role in determining the influence of such information on users' subsequent behavior. We find that under certain conditions, users become less susceptible to ambient social information even in online platforms that provide such information in abundance. Thus, we show that the role of ambient social information in shaping behavior is more complicated than previously believed.

Chapter 3 Whom Do They Prefer for Making Money? Determinants of Information Source Preference in Online Platforms

1. INTRODUCTION

In the recent years, the emergence of online platforms has created newer ways for information flows. For example, on Facebook, over 2 billion individuals generate information about their daily activities, which is accessible to their network of friends and even beyond (https://techcrunch.com/2017/06/27/facebook-2-billion-users/). On YouTube, individuals collectively view nearly 1 billion hours of videos per day (https://techcrunch.com/2017/02/28/people-now-watch-1-billion-hours-of-youtube-per-day/), observing information about diverse topics, including the video uploaders' daily activities, product reviews, and opinions about current affairs. In sum, online platforms have led to a considerable increase in the quantity of information produced, as well as the ease with which it can be accessed. We consider such information to be social because it conveys users' prior actions and choices (Shang & Croson, 2009) as well as ambient because it is enveloping and ever-present (Leonardi, 2015).

As the existing research reveals, ambient social information is highly consequential in shaping the audience's subsequent behavior (Bapna & Umyarov, 2015; Aral & Walker, 2011). However, users' capacity to consume such information remains limited. Therefore, a focal user may be selective in that she may prefer some information sources more than the others. In this regard, our fundamental research question is as follows: with several users generating ambient social information and acting as information sources in online platforms, whom does the focal user prefer? This question is even more pertinent when the preference for a particular information source has considerable financial implications.

To address this question, we begin with an observation that the preference for a specific information source is the function of different attributes of the source. The focal user can observe these attributes and then formulate her preference. However, different attributes affect the focal user's preference in distinct ways. On the one hand, certain attributes about the information source may have a positive effect on preference. For instance, studies of herding signal show that the focal user may prefer those who have already attracted a large audience (Liu et al., 2015). Similarly, the literature also suggests that the prior payoffs that the focal user has received from an information source also increase that user's preference for that source (Kaustia & Knüpfer, 2008). However, these studies fail to incorporate the counteracting effect of physical distance between the information source and focal user. The literature indicates that physical distance impedes online interpersonal exchanges, including those of a financial nature (Agarwal, Catalini, & Goldfarb, 2015). Building on this work, we examine *how the combined effect of physical distance, herding signal, and prior payoffs drive the focal user's preference for an information source.*

Our empirical context is an online social trading platform. Such platforms are a part of recent technological innovations that are changing the way financial markets operate (Doering, Neumann, & Paul, 2015). In the coming years, social trading and other financial technologies, which are often referred to as Fintech, are expected to gain importance as millennials account for an increasingly large section of the active investor population (Morris, 2016). For several reasons, it is appropriate to use a social trading platform as an empirical context for studying the research question at hand. First, the platform continuously records and displays investment transactions of each user. Thus, at any time, multiple information sources are present. Second, social trading platforms incorporate a novel feature through which one can unambiguously observe variations in users' preferences. This is made possible because a focal user's preference for an information source is indicated by the exchange of funds from the former to the latter. Third, the platform has a global user-base. In our dataset, we observe users from 95 distinct countries. Thus, the physical distance between the user and the information source may vary greatly, affecting the former's preference for the latter. Finally, because preferring one information source over the other has direct implications for the focal user's investment portfolio. Hence, for the focal user, being selective in preference for different information sources is a meaningful outcome.

Our study formulates and tests a set of related hypotheses, each proposing a link between an attribute of an information source and the focal user's preference for that source. Our estimation is based on a dataset comprised of forty thousand user-source dyads between actively trading users, observed for 46 consecutive weeks. Our results indicate that while the herding signal and prior payoffs increase the focal user's preference for an information source, the physical distance between the two moderates these effects. In particular, *the herding signal of a distant information source has a significantly stronger effect than that of a proximate information source*, suggesting that while the herding signal as a decision heuristic may positively affect the focal user's preference, its effect is more meaningful in the presence of physical distance (Agarwal, Catalini, & Goldfarb, 2015).

We also find that the focal user significantly discounts the prior payoffs obtained from a physically distant information source compared to those obtained from a physically proximate information source, indicating that the unfamiliarity which physical distance breeds may supplant the effect of the focal user's prior payoffs from an information source (Kaustia & Knüpfer, 2008). This finding strongly points to a behavioral explanation of the physical distance (Lin & Viswanathan, 2015). Collectively, our study provides key insights into how, in the presence of multiple information sources, the attributes of a specific information source are factored into the focal user's preference for that source.

The rest of the article is structured as follows: in the next section, we review the relevant literature and describe the study's empirical context. Next, we develop a set of related hypotheses. In the next section, we provide variable description and the dataset construction process, followed by results of hypothesis testing. We complete the analysis by demonstrating the robustness of our findings. In the final section, we discuss the contributions of the study and identify directions for future research.

2. RELATED WORK

2.1 Attributes of information sources that explain focal user's preference

A considerable body of work suggests that users' behavior in online platforms can be explained by the ambient social information generated by the multitude of platform users, who act as information sources (Aral & Walker, 2012, Chen, Wang, & Xie, 2011). However, a more granular argument suggests that all information sources do not matter equally (Kim & Ratchford, 2012). Constrained by her ability to consume the available information, a focal user prefers some information sources over others. Studies have attributed the uneven preference to the attributes of the information source. Such attributes range from simple demographic features to more nuanced psychological characteristics such as identity.

Aral & Walker (2012) showed that on Facebook, product adoption notification from older, single males generate greater influence over other users than the notifications generated by users from other demographic clusters. While in their randomized field experiment, authors do not model the focal user's preference as a deliberate choice, the *demography-based variations* are indicative that all information sources are not considered equal in online settings. Other studies examine the presence of *prior friendship ties* between the focal user and information sources as another determinant of uneven preference. Shi & Whinston (2013) show that the check-in information provided by the focal user's friends has a strong influence on the likelihood of that user's subsequent visit to that location because friends' ambient social information acts as a strong signal of "congruence in preferences." Similarly, Liu et al.(2015) show that when contributing to a crowdfunding project, the focal user is more likely to follow the investment actions of her friends. Lastly, the focal user may prefer an information source whose *identity* is known. Forman, Ghose, & Wiesenfeld (2008) show that a focal user perceives those online product reviews, in which the reviewer's identity is revealed, as more helpful than those in which the reviewer identity remains hidden. In sum, a considerable body of existing work identifies several attributes of an information source, which may increase the focal user's preference for that source.

On the other hand, some attributes may lower the focal user's preference for an information source. Recent literature on physical distance effects in online platforms, especially in Crowdfunding platforms, shows that individuals prefer to donate or invest money in physically proximal campaigns. There exist several explanations for the distance effect in online platforms. For instance, Lin & Viswanathan (2015) find behavioral factors drive the lead individuals to invest in proximal campaigns even when such campaigns provide lower economic returns (Lin & Viswanathan, 2015). Burtch, Ghose, & Wattal (2014) indicate that higher physical distance operates as a proxy for greater unfamiliarity and therefore, inhibits interpersonal interactions and exchanges. In closing, physical distance acts as an inhibitor for interpersonal exchanges in online platforms and therefore may counterbalance the effects of the other attributes of an information source.

2.2 The combined effect of different attributes of the information source

Based on the prior work, we argue that the complete understanding of which information source users prefer cannot be obtained unless one simultaneously examines these counteracting effects. So far, the literature remains unclear on how physical distance combines with other attributes of the information source. For example, in an e-commerce setting, Forman et al. (2008) show that the physical distance between the information source and the users weakens the effect of additional, identity-revealing information about the source. On the contrary, in online crowdfunding settings, Agarwal, Catalini, & Goldfarb (2015) show that the physical distance between the entrepreneur and the investor may enhance the effect of the entrepreneur's herding signal. Lastly, a similar study by Greenwood & Gopal (2017) in the context of offline IT investments finds no consistent empirical support for the interaction between the IT entrepreneur's herding signal and their physical distance from the VC.

Together, these studies underscore that present literature is highly inconclusive in ascertaining how the focal user's physical distance from an information source is factored into the evaluation of the source's other attributes. The present study addresses this research gap. In particular, we model the focal user's preference for an information source as a function of herding signal, prior payoffs and physical distance from that source.

2.3 User behavior in online platforms for financial markets

The present study also relates to the growing interest in examining human behavior in emerging online platforms for financial markets. Collectively referred to as Fintech, these platforms have become increasingly popular in research as well as practice. For instance, Crowdfunding researchers have examined the antecedents and consequences of prosocial donations and contributions to public good (Burtch, Ghose, & Wattal, 2013; Liu et al., 2015), investments in real estate ventures, and funding given to different forms of creative arts such as music and films (Gamble, Brennen, & McAdam, 2017; Agarwal, Catalini, & Goldfarb, 2015). Another prominent technological advance under Fintech refers to cryptocurrencies. For example, recent studies of Bitcoin user networks find predictable patterns such as preferential attachment (Kondor et al., 2014) and trust (Sas & Khairuddin, 2017). Lastly, researchers have also examined behavior in social trading platforms. Prominent research efforts in this stream include prediction of financial returns through trades (Berger, Wenzel, & Wohlgemuth, 2017) and the disposition effect, which refers to realizing investment returns by immediately closing a profitable investment while retaining a losing investment for a longer time (Glaser & Risius, 2017). The present study is motivated by such attempts to explain behavior in information-rich, Fintech settings. More specifically, we model the focal user's preference for investment signals generated by other traders in a social trading platform (Wohlgemuth, Berger, & Wenzel, 2016).

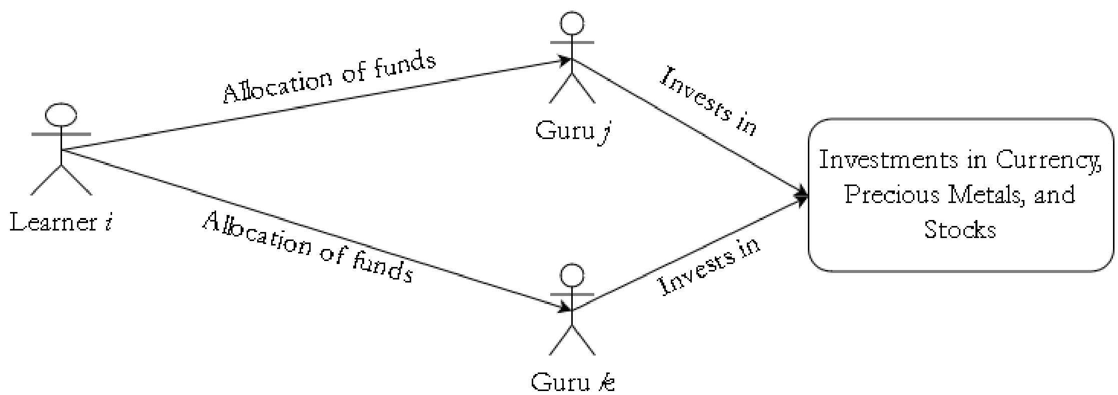
3. STUDY CONTEXT

In this section, we describe the empirical context of the study to clarify the link between the discussion in the previous section and the subsequent development of hypotheses. The study is based on an online social trading platform, which we refer to as XTrader (a pseudonym). A typical user on this platform is a trader who is interested in investing money in several instruments such as company stocks, precious metals, and currencies. Social trading platforms apply features from a typical social media platform to online trading (Doering, Neumann, & Paul, 2015). For instance, social trading platforms allow traders to communicate with each other, share updates about their trading activities, and form ties.

The study leverages a salient feature common across many social trading platforms. It is referred to as copy-trading (Figure 3.1). Essentially, it is a mechanism that creates directed trader- dyads. The feature works as follows: a trader i joins the platform by depositing a certain fund amount (f_i), which she wants to invest. This funding amount can be deposited with the trader's profile using a variety of means such as Bank account, and credit cards. Once on the platform, i can observe ambient social information, comprised of other traders' investment signals (Appendix II). Based on this information, trader i can choose another trader j and allocate a portion of f_i to trader j. Upon allocation, the platform automatically replicates all of trader j's subsequent and currently open investments to trader i's portfolio.

Thus, through the allocation of funds, trader *i*' indicates her preference for trader *j* as an information source. Moreover, because trader *i* has limited funds, she has to choose *j* from a number of other traders deliberately. Therefore, non-zero allocation in a trader dyad *{i* → *j}* indicates that trader *i* explicitly picked trader *j* from a sizeable pool of other traders, making trader *i*'s preference for a trader *j*, a deliberate action. In the rest of the study, the allocating trader *i* is referred to as Learner while trader *j* is referred to as the Guru.

Figure 3.1: Copy-Trading Mechanism in Social Trading Platforms



4. HYPOTHESIS DEVELOPMENT

4.1 Herding signal of an information source

The study's first main effect pertains to the herding signal. Herding has been widely studied in offline (Sun, 2013) as well as online settings, including in settings involving financial and investment decisions (Agarwal, Catalini, & Goldfarb, 2015). It is said to occur when a decision maker imitates the prior, collective behavior of others. The existing literature identifies several explanations for herding behavior, such as social proof (Rao, Greve, & Davis, 2001), information cascades (Duan, Gu, & Whinston, 2009), and observational learning (Shi & Whinston, 2013). However, across all these explanations, the basic herding mechanisms remain the same. By observing the prior behavior of others, individuals follow the herd (Greenwood & Gopal, 2017). Extending the same argument, we postulate our first hypothesis:

<u>Hypothesis-1:</u> A herding signal of an information source will increase the focal user's preference for that source.

4.2 Prior payoffs from an information source

In the second hypothesis, we postulate the link between the focal user's prior payoffs from an information source and her subsequent preference for that source. The literature suggests that individuals "tend to repeat actions that have produced favorable outcomes in the past" (Kaustia & Knüpfer, 2008; p. 2683) as positive payoffs from her prior actions are likely to result in choice reinforcement. This hypothesis has found considerable empirical and theoretical support, including in investment decisions (Kaustia & Knüpfer, 2008). In the same vein, we argue that if the focal user obtains positive, prior payoffs from a particular information source, the user is likely to exhibit a higher preference for that source in the subsequent time. Thus, our second hypothesis is as follows:

<u>Hypothesis-2:</u> Prior payoffs generated by the focal user from an information source will increase the focal user's preference for that source.

4.3 The physical distance between the focal user and the information source

Having hypothesized the effects of attributes that increase the focal user's preference for an information source, we proceed to examine the counteracting effect of physical distance between the focal user and an information source. With the emergence of the internet, it was opined that “distance is dead” (Cairncross, 2001) as well as “the world has become flat” (Friedman, 2005) because individuals face no additional costs for interacting with those who are physically distant. However, several studies have shown that physical distance still matters. In this regard, several mechanisms have been presented, explaining the role of physical distance in online settings. These include home-bias, which is particularly applicable for goods that can be consumed only in a specific location (Hortacsu, Martinez-Jerez, & Douglas, 2009), absence of prior social ties with physically distant users (Agarwal, Catalini, & Goldfarb, 2015), and distance as a proxy for unfamiliarity and lack of awareness (Burtch, Ghose, & Wattal, 2014). However, irrespective of the mechanism, the result remains consistent. The greater the physical distance between two users, the less likely are they to interact and participate in an exchange. Extrapolating the same logic, we formulate the third hypothesis as follows:

Hypothesis 3: A focal user will have a lower preference for a physically distant information source than for a physically proximal information source

4.4 The interaction between herding signal and physical distance

Having postulated the three main effects, we move on to suggesting ways in which the physical distance effect may interact with those of herding signal and prior payoffs. The herding signal and physical distance, as topics, are covered in distinct streams of literature. Therefore, little is known about the potential interaction between the two attributes. While some recent studies have begun exploring this issue, especially in financial investment contexts, the evidence remains contradictory. In their study of venture capitalist (VC) investments in IT companies, Greenwood & Gopal (2017) hypothesize that the herding signal will have a stronger effect on VC’s investment likelihood as the physical distance between the entrepreneur and the VC increases. However, they find no consistent empirical support for this claim (p. 1002). On the other hand, Agarwal, Catalini, & Goldfarb (2015) find that in online crowdfunding, an investor is significantly more influenced by the herding signal of a distant entrepreneur than that of a physically proximal entrepreneur.

In the present study, we hypothesize that the herding signal of a physically distant information source will have a stronger effect than that of a physically proximal source. Our argument borrows from the literature on the role of physical distance in financial investment decisions (Huberman, 2001). As suggested in H3, users may perceive a physically proximal source as more familiar than a physically distant one. We argue that such preexisting sense of familiarity will override the source's herding signal as a decision heuristic, thereby reducing its influence on the focal user's preference for an information source. Conversely, the herding signal may be significantly more important when the focal user and the information source are physically distant because the distance implies greater unfamiliarity (Burtch, Ghose, & Wattal, 2014). Thus, our penultimate hypothesis is as follows:

<u>Hypothesis 4:</u> Physical distance between the focal user and the information source will increase the effect of the source's herding signal on the focal user's preference for that source.

4.5 The interaction between prior payoffs and physical distance

Lastly, we hypothesize an interaction effect between the focal user's prior payoffs obtained from an information source and the physical distance. The effect may operate in either way. On the one hand, physical distance, and the unfamiliarity it espouses may weaken the positive effect of prior payoffs. As a result, the focal user may not perceive the payoffs obtained from a physically distant information source to be evidence of benefits of preferring that source over the rest. In other words, physical distance may act as a boundary condition for the extent to which past outcomes drive future actions (Kaustia & Knüpfer, 2008).

On the other hand, prior payoffs, especially those obtained in monetary terms, may overcome the unfamiliarity barrier associated with physical distance. In such a scenario, the payoffs may demonstrate the viability of preferring an information source, to the extent that the focal user may at least partially disregard the unfamiliarity concerns related to physical distance. Because the extant literature provides no clear answer to the directionality of the interaction effect between physical distance and prior payoffs, we postulate our final hypothesis in two parts.

<u>Hypothesis 5a:</u> The greater the physical distance between the focal user and the information source, the weaker will be the positive effect of prior payoffs on the focal user's preference for that information source.

<u>Hypothesis 5b:</u> The greater the physical distance between the focal user and the information source, the stronger will be the positive effect of prior payoffs on the focal user's preference for that information source.

5. DATASET AND VARIABLE DESCRIPTION

5.1 Dataset construction

Our initial dataset was comprised of 57,324 distinct trader dyads that existed at least once during the observation window of 46 weeks. To construct the final dataset, we first combine every dyad *{i → j}* with every week *t*. Thus, we construct 57,324 balanced panels for 46 weeks, obtaining a total of 2,636,904 observations. From this dyad, we exclude all observations in which a Learner's collective allocation crossed the 100 percent threshold. Such observations may result from potential measurement errors. Thus, our observation set reduces to 2,423,451. The number of dyads remains constant at 57,324. Next, we consider the possibility that observing every dyad in every week may not be valid as there may exist unobserved, time-varying heterogeneity across trader dyads, which may influence the value of the outcome variable. To overcome any such concern, we enforce the following baseline constraint. We only include a dyad *{i → j}* in week *t* if at least one trade was posted to the accounts of trader *i* and trader *j*, which suggests that both the traders in a dyad were not only present on the platform but also were actively investing. This constraint reasonably assures that the values of the outcome variable indeed stem from deliberate decisions of active traders. Implementing such constraints is important in studies using observational data (Chua, Roth, & Lemoine, 2015). After applying this condition, the consideration set reduces to 126,120 observations, comprised of 42,105 distinct trader pairs. We use this dataset for estimation.

5.2 Panel data structure

The dataset consists of repeated observations of trader dyads over time. The panel structure can be created in many ways. For instance, we can set up a panel for every trader dyad observed over the data collection window. However, doing so may introduce error correlations between panels and hence, create simultaneity. For instance, consider the following example: trader *i* who allocates to both trader *j* and trader *k* in week *t.* Thus, setting different for every trader dyad will result in the following structure:

$$\text{Panel}_{\{i,j\}} \begin{cases} \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t\text{-}k} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t+k} \end{cases}$$

Thus, the values of the outcome variable in any subsequent panel {i,k}, which comprises of Learner *I* and another Guru *k* in the same week *t* will be dependent on the values of the outcome variable in panel {i,j} observed in the same week because Learner *i* allocates to different Gurus from the same fund. Thus, allocation to one Guru will simultenously constrain the Learner's ability to allocate to other Gurus at the same time. In other words, there will be a dependency between the panels of any given Learner.

To overcome this issue, we construct panels for each Learner *i* that includes all the Gurus she allocated funds to during the observation window and cluster the standard errors on *i*. Thus, the number of panels in the consideration set is equal to the number of distinct Learners. Thus, the dataset, which is retains its dyadic nature, is structured as follows. Note that the panel variable is Learner *{i}*.

$$\text{Panel}_{\{i\}} \begin{cases} \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t\text{-}k} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{j}\}_{\text{week } t+k} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{k}\}_{\text{week } t} \\ \{\text{Learner } \boldsymbol{i} \rightarrow \text{Guru } \boldsymbol{k}\}_{\text{week } t+k} \end{cases}$$

Because all the observations of Learner *i* are captured within a single panel, we reasonably overcome the issue of dependency between panels. As explained later, for estimation, we incorporate dummies for Gurus and for Observation weeks to account for any unobserved heterogeneity for these entities. Lastly, this approach allows us to estimate the effect of physical distance, which is time-invariant in a given dyad. This form of panel data structure is typical in dyadic studies involving distance-based arguments (Burtch, Ghose, & Wattal, 2014).

For the hypothesis testing, we use a linear estimator devised particularly for large datasets requiring multi-way fixed effects specification (Correia, 2016; Torres et al., 2013). The principal advantage of this estimator is its computational efficiency, which is essential when the estimation involves large vectors of dummies. The estimation incorporates three fixed effects. First, we control for any unobserved, time-invariant heterogeneity across the Learners. We compute heteroskedastic-robust standard errors, clustered on Learners. Next, we also incorporate dummy vectors for the Guru as well as the observation week. Thus, our estimates are robust to the potentially confounding, time-invariant trader attributes such as gender, which has a bearing on investment behavior (Sundén & Surette, 1998) as well as to any exogenous time shocks.

5.3 Variable descriptions

<u>Outcome variable:</u> the study's outcome variable is the percentage of funds allocated by a Learner to a Guru in a given week (ALLOC). In the dataset, we observe a considerable variation across the allocation percentages, ranging from 0 to 100. In approximately 43.9% of the observations, we find that a Learner does not allocate any funds to the Guru while in roughly 8.3% of observations, a Learner allocated 100% of her funds to the Guru. In the remaining observations (47.7%), the allocation percentage is between 0 and 100.

Independent variables: To construct the measures for Learner's prior payoffs from a Guru (PRIOR_PAYOFFS), we draw on the transaction-level information provided by the platform. Specifically, we cumulate the gain across all the transactions, which were posted to Learner *i*'s portfolio in week *t* but were initiated by Guru *j*. The collective gain of such transactions represents the benefit that Learner i realized by choosing Guru *j* in week *t*. Wherever the collective gain from Guru *j* to Learner *i* was missing, we set it to 0, indicating that *i* generated no benefit from *j* in that week. For the estimation, we incorporate a 1-week lag to avoid simultaneity between the outcome variable and the Learner's prior payoffs from allocations to that Guru. Our measure of prior payoffs is consistent with those used in the literature. For instance, Borah & Tellis (2014) measure prior payoffs of similar, past actions (p. 123).

The study's second explanatory measure pertains to the herding signal (COPIER_COUNT). We calculate this variable as follows: for each Guru *j*, we calculate the total number of Learners in week *t*. From this count, we subtract 1 to ensure that the focal Learner *i* is not counted. The resulting count of Learners is used as a herding signal. For instance, consider Learners i_1 and i_2 who copy-traded Guru *j* in week *t*. Thus, for the dyad $\{i_1 \rightarrow j\}$, the value of Guru *j*'s herding signal is 1. Such count-based measures are typical in measuring the herding effect (Duan, Gu, & Whinston, 2009; Sun, 2013).

The study's final explanatory variable is the physical distance. We adopt Mayer & Zignago's (2011) geodesic distance computed using the Great Circle formula. The measure calculates the distance between the two most populous cities of a given country pair (Mayer & Zignago, 2011; p. 10). In our sample, we observe that the traders are spread across 95 unique countries. Thus, the dataset captures a globally dispersed sample of users, underscoring the pertinence of physical distance as a factor. Our choice of measure for physical distance is consistent with the recent studies of distance effects in online contexts (Burtch, Ghose, & Wattal, 2014).

Control Variables: We incorporate several control variables to tease out the hypothesized effects. To account for any unobserved, time-invariant confounders, we incorporate fixed effects for Learners and Gurus. Next, to capture unobserved, temporal shocks, we incorporate a vector of dummies for the 46 observation weeks. Further, we also include a number of time-varying, Learner-specific controls.

First, we control for the Learner's as well as Guru's tenures on the platform, which we calculate as the count of days between the date on which the trader joined the platform and the first date of the given observation week. Note that Guru and Learner may have joined the platform before the data collection and hence, the 46-week observation window does not constrain calculation of tenure. Second, we control for the number of Gurus a given Learner copy-traded in the prior week (PRIOR_GURU_COUNT). Because the total fund amount limits the Learner, we expect the allocation to a specific Guru to be lower if the Learner is copy-trading with multiple Gurus. We also include the count of weeks in which the Learner profited through self-trading (SELF_TRADE_BENEFITS). Collectively, such instances are likely to nudge a Learner to trade independently and therefore, have lower allocation.

Lastly, we control for the possible audience effect that could influence the Learner's copy-trading behavior. On XTrader, the trading activities and portfolios of the traders are publicly visible. Given this information transparency, a Learner may be less inclined to portray herself as relying on Guru's investment signals (Lee, 2002), especially when other traders are observing her behavior. Therefore, the inhibiting effect of "watching eyes" may lower the Learner's allocation. This effect alludes to recent studies of behavior in online platforms (Huang, Hong, & Burtch, 2016). We control for the potential effect of "watching eyes" by incorporating the number of distinct traders who have communicated with a focal Learner (MESSAGE_SENDERS). For estimation, we log-transform the control variables. Table 3.1 summarizes the key variables and lists the data sources.

Table 3.1: Variable Description and Data Source

Measure **[Role in the model]**	**Description**	**Source**
ALLOC [Outcome]	The percent allocation from Learner to Guru in week *t*	The platform
LEARER_TENURE [Control]	The days between the date on which the Learner joined the platforms and the first date of the given week	The platform
PRIOR_GURU_COUNT [Control]	The count of Gurus, which a Learner copied in a week	The platform
SELF_TRADE_BENEFITS [Control]	The cumulative count of Learner's profitable weeks through self-trading during the observation window	The platform
MESSAGE_SENDERS [Control]	The cumulative count of distinct traders who sent messages to the Learner	The platform
GURU_TENURE [Control]	The days between the date on which the Guru joined the platforms and the first date of the given week	The platform
PRIOR_PAYOFFS [Explanatory]	The aggregate gain for Learner *i* through copied transactions from Guru *j* in week *t-1*	The platform
COPIER_COUNT [Explanatory]	Count of Learners copying a given Guru in a given week *t-1*	The platform

PHYSICAL_DISTANCE [Explanatory]	The physical distance (in KMs.) between the two largest cities in Learner's and Guru's respective home countries	CEPII We use "distcap" variable from dist_cepii.dta

5.4 Estimation strategy

To test the hypotheses, we employ the following model specification:

$$Y_{ijt} = \beta_0 + \beta_1 \text{ Gains from Guru} + \beta_2 \text{ Count of other Learners} + \beta_3 \text{ Physical Distance} + \beta_4(\text{Physical Distance} \times \text{Gain from Guru}) + \beta_5(\text{Physical Distance} \times \text{Count of other Learners}) + \mu_{it} + \sigma_{jt} + \phi_i + \delta_j + \lambda_t + \varepsilon_{ijt} \quad \textbf{(Equation 1)}$$

The left-hand side of equation 1 is the allocation from Learner *i* to Guru *j* in the observation week *t*. The right-hand side depicts the linear combination of explanatory as well as the control variables. In addition to the coefficients of interest ($\beta_{1\text{-}5}$), we also incorporate time-varying controls for the Learner (μ_{it}) and the Guru (σ_{jt}), Learner fixed effects (ϕ_i), Guru fixed effects (δ_j), and the observation week fixed effects (λ_t) As per H1 and H2, the coefficient β_1 and β_2 should be significant and positive. According to H3, β_3 should be significant and negative. Next, we have developed competing hypotheses, indicating that while both β_4 and β_5 should be significant, each coefficient can attain either positive or negative sign. Finally, we include an error term, ε_{ijt}. Some scholars recommended using the fractional response estimator for percentage-based outcomes because it is difficult to ensure the bounded nature of the outcome variable (Papke & Wooldridge, 2008). However, non-linear estimators often suffer from the incidental parameter problem (IPP) when the estimation involves multi-way fixed effects. Therefore, we use a linear estimator with multi-way fixed effects (Correia, 2016). To address the concern about out-of-bounds predictions, we report the proportion of the projected values of the outcome variable, (ALLOC), which fall outside the [0,1] interval.

In less than 1% of the observations, we find such values. Therefore, we conclude that our choice of estimator is reasonably justified.

6. ANALYSIS

6.1 Descriptive statistics and collinearity diagnostics

Before undertaking the formal estimation, we check the dataset for indicators of possible multicollinearity. We begin by reporting the pairwise correlations between the variables (Table 3.2) followed by the descriptive statistics (Table 3.3). We find that all the correlation coefficients, while significant, are less than 0.5. Also, the highest variance inflation factor (VIF) is 1.29, which is well below the expected threshold (Rai et al., 2015). Therefore, we conclude that the variables exhibit no major collinearity concerns.

Table 3.2: Pairwise Correlations and Collinearity Diagnostics

	Allocation	Learner Exp (Log)	Guru Exp (Log)	Learner's Count of Gurus (Log)	Count of Profitable Week (Log)	Count of Message Senders (Log)	Count of Copiers	Gain in Pair	Geographic Distance
Allocation	1								
Learner Exp (Log)	-0.0695*	1							
Guru Exp (Log)	0.1005*	0.1948*	1						
Learner's Count of Gurus (Log)	-0.1264*	0.0415*	0.0124*	1					
Count of Profitable Week (Log)	-0.1603*	0.1647*	-0.0348*	-0.0703*	1				
Count of Message Senders (Log)	-0.1751*	0.2135*	-0.0845*	-0.0353*	0.3981*	1			
Count of Copiers	0.1388*	0.0865*	0.4282*	0.0108*	-0.0319*	-0.0833*	1		
Gain in Pair	0.1001*	-0.0185*	-0.0139*	-0.0044	-0.0135*	-0.0141*	-0.0007	1	
Geographic Distance	0.0024	-0.0248*	-0.0063*	-0.0026	-0.0213*	-0.0211*	-0.0557*	0.0078*	1

* $p < 0.05$
Max VIF = 1.29
Average VIF = 1.14

Table 3.3: Descriptive Statistics

	Allocation	Learner Exp (Log)	Guru Exp (Log)	Learner's Count of Gurus (Log)	Count of Profitable Week (Log)	Count of Message Senders (Log)	Count of Copiers	Gain in Pair	Geographic Distance (Normalized)
Mean	19.82658	6.18069	6.514077	1.167455	0.336896	0.68878	443.5731	6.401245	-0.10797
Minimum	0	0	0	0	0	0	0	-6902.26	-1.06843
Maximum	100	7.723562	7.659643	3.044523	2.772589	8.540323	2119	2594.9	3.271918
Standard Deviation	29.88819	0.617133	0.60571	0.8389883	0.5344	1.209697	578.861	89.64584	1.00282

6.2 Main analysis

We begin by estimating the influence of the Learner-specific, time-varying control variables on the allocation percentages (Model 1). The control variables operate in the expected fashion. For instance, SELF_TRADE_BENEFITS up to the prior week translate into lower allocation in the current week towards a specific Guru, suggesting that the Learners who obtained gains through self-trades allocate a smaller portion of their funds. Similarly, the GURU_COUNT in the prior week corresponds with a lower percentage allocation in the current week in a dyad; that is, the higher the number of Gurus a Learner copy-traded with, the more sparsely is her fund distributed.

Next, in Model 2, we introduce the HERDING_SIGNAL as the first explanatory variable. In Model 3, we include PRIOR_PAYOFFS from the preceding week. Once again, we obtain a significant and positive coefficient, which is consistent with the argument in H2. In Model 4, we estimate the last main effect by including the DISTANCE measure. In Model 5, we find that the effects postulated in H1-H3 are collectively significant. Coming to the effect sizes, we find that a unit increase in the HERDING_SIGNAL results in 0.0198 units increase in allocation while a similar increase in PRIOR_PAYOFFS increases the allocation by 0.0192. On the other hand, a unit increase in PHYSICAL_DISTANCE is associated with a reduction in allocation by 0.54 units.

Moving on to estimating the interaction effects, we find that the interaction between HERDING_SIGNAL and PHYSICAL_DISTANCE is positive (Model 6). Thus, H4 is supported. In Figure 3.2, we observe that higher HERDING_SIGNAL marginally mitigates the negative influence of PHYSICAL_DISTANCE. Next, in Model 7, we estimate the interaction between PRIOR_PAYOFFS and PHYSICAL_DISTANCE, finding support for H5a. Figure 3.3 reveals that PHYSICAL_DISTANCE attenuates the positive effect of PRIOR_PAYOFFS. In the fully specified Model 8, we find that the estimates are jointly significant. Thus, we conclude that the hypotheses H1-H5a are supported (Table 3.4). Lastly, to test whether the effects are stable to the inclusion of cultural distance, which are known to affect investment behavior (Beckmann, Menkhoff, & Suto, 2008), we re-estimate Model 8 by including Hofstede's dimensions of culture: *power distance, masculinity, individualism,* and *uncertainty avoidance* (https://www.hofstede-insights.com/product/compare-countries/). Including the absolute difference between these scores for countries of traders in a given dyad, we find coefficients similar to those obtained in Model 8.

Figure 3.3 also indicates an interesting pattern. For Gurus who generated very low prior payoffs, physical distance seems to mitigate the negative effect on allocation. In other words, a Learner has a higher preference for a low-performing yet distant source over a comparably low-performing but proximal source. A possible explanation is that for a Learner, evaluating a distant source's investment strategies based on prior payoffs may be relatively more obscure than those of a proximal source. This effect further underscores the behavioral nature of the physical distance effect.

Figure 3.2: Interaction between PHYSICAL_DISTANCE and HERDING_SIGNAL

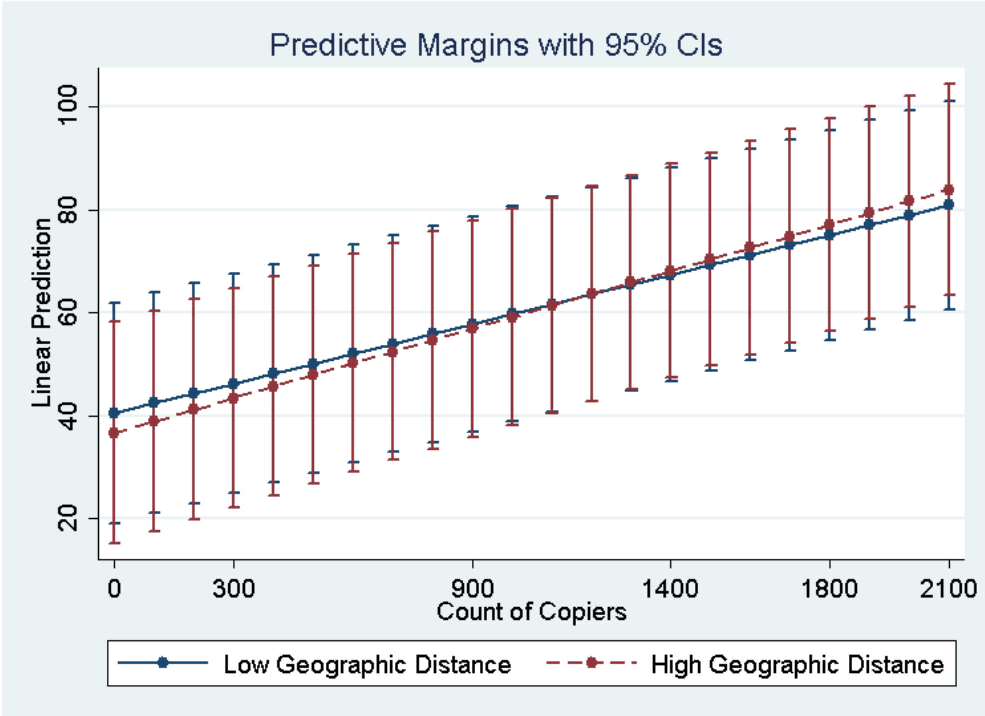


Figure 3.3: Interaction between PHYSICAL_DISTANCE and PRIOR_PAYOFFS

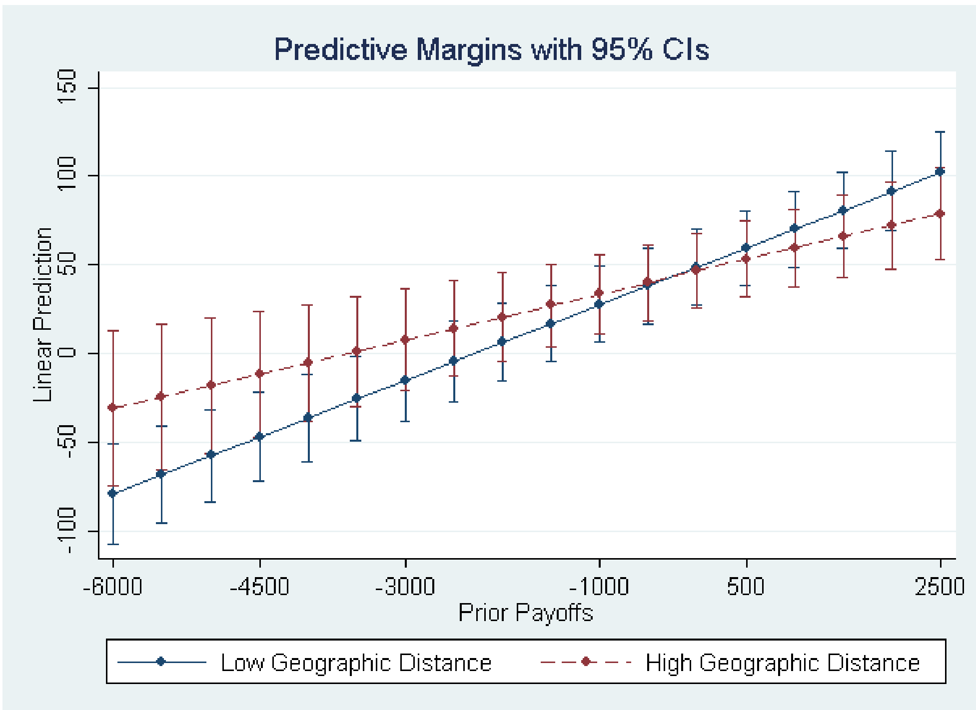


Table 3.4: Main Analysis with REGHDFE

(DV = Allocation in a Trader Dyad)

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Learner Exp (Log)	0.256	0.738	0.267	0.272	0.769	0.731	0.780	0.742
	(0.977)	(0.976)	(0.969)	(0.977)	(0.968)	(0.968)	(0.968)	(0.968)
Guru Exp (Log)	35.35***	6.854***	34.42***	35.34***	5.591***	5.643***	5.554***	5.606***
	(1.430)	(1.445)	(1.405)	(1.430)	(1.429)	(1.431)	(1.429)	(1.431)
Learner's Count of Gurus (Log)	0.0895	0.0375	0.0640	0.0901	0.0114	0.0125	0.00959	0.0107
	(0.0866)	(0.0857)	(0.0867)	(0.0866)	(0.0860)	(0.0860)	(0.0860)	(0.0860)
Count of Profitable Week (Log)	-1.124**	-1.485***	-1.009**	-1.124**	-1.372***	-1.378***	-1.366***	-1.373***
	(0.343)	(0.341)	(0.340)	(0.343)	(0.339)	(0.339)	(0.339)	(0.339)
Count of Message Senders (Log)	-1.987***	-2.221***	-1.926***	-1.988***	-2.162***	-2.164***	-2.162***	-2.163***
	(0.174)	(0.173)	(0.173)	(0.174)	(0.171)	(0.171)	(0.171)	(0.171)
Count of Copiers		0.0196***			0.0198***	0.0200***	0.0198***	0.0200***
		(0.00065)			(0.00064)	(0.00065)	(0.00064)	(0.00065)
Gain in Pair			0.0188***		0.0192***	0.0193***	0.0192***	0.0192***
			(0.00117)		(0.00118)	(0.00118)	(0.00116)	(0.00116)
Geographic Distance				-0.543*	-0.546*	-0.952***	-0.538*	-0.942***
				(0.231)	(0.229)	(0.246)	(0.229)	(0.246)
Count of Copiers×Geographic Distance						0.000821***		0.000818***
						(0.00024)		(0.00024)
Gain in Pair×Geographic Distance							-0.00212*	-0.00208*
							(0.00102)	(0.00103)
N	95227	95227	95227	95227	95227	95227	95227	95227
Adj. R^2	0.0217	0.0385	0.0288	0.0218	0.0462	0.0467	0.0463	0.0468
AIC	807132.3	805478.7	806437.9	807115.8	804720.4	804675.2	804711.3	804666.4
BIC	807179.6	805535.5	806494.7	807172.6	804796.1	804760.3	804796.5	804761.1
Learner Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Guru Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y

Robust standard errors clustered on Learners are reported in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.3 Robustness check I: Are the results dependent on panels in which the allocation does not change during the observation window

Similar to the traditional OLS estimator with fixed effects, the linear estimator with multi-way fixed effects, which is used in the present study (Correia, 2016) retains a panel even if the outcome measure exhibits no variation during the entire observation window. However, such panels may be systematically different from the rest because the changes in the explanatory variables are not reflected in the outcome variable. As the next robustness check, we assess whether such panels are driving the results. We begin by calculating the standard deviation of the outcome measure for each Learner. Then, we exclude the panels of those Learners in which the standard deviation was zero, suggesting that for the given Learner, the outcome measure has not changed during the 46-week long window. In all, we find 20,754 such pairs, amounting to 44,978 observations. After excluding these panels, we obtain qualitatively similar results (Model 9). Therefore, we claim that estimates are robust to the presence of Learners in which the outcome variable does not change (Table 3.5).

Table 3.5: Robustness Checks

(DV = Allocation in a Trader Dyad)

	Model 9
Learner Exp (Log)	1.219
	(1.316)
Guru Exp (Log)	10.05***
	(1.927)
Learner's Count of Gurus (Log)	-0.441***
	(0.116)
Count of Profitable Week (Log)	-1.352**
	(0.479)
Count of Message Senders (Log)	-2.279***
	(0.238)
Count of Copiers	0.0211***
	(0.000797)
Gain in Pair	0.0137***
	(0.00109)
Geographic Distance	-0.680*
	(0.288)
Count of Copiers×Geographic Distance	0.000646*
	(0.000257)
Gain in Pair×Geographic Distance	-0.00332***
	(0.000860)
N	63046
Adj. R^2	0.0587
AIC	531570.1
BIC	531660.7
Learner Fixed Effects	Y
Guru Fixed Effects	Y
Observation Week Fixed Effects	Y

Robust standard errors clustered on Learners are reported in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.4 Robustness check II: Are the results robust to the changes in the underlying observation set?

Next, we check whether changing the structure of the underlying dataset changes the results. This exercise assesses the stability of the results to smaller sample sizes (Lin, Lucas, & Shmueli, 2013; Yao, Dresner, & Palmer, 2009). Accordingly, we extract random subsamples, by excluding 10%, 20%, 30%, 40% and 50% of the observations.

For each subsample, we obtain consistent estimates (Model 10-14), suggesting that results are robust to the sample size (Table 3.6).

Table 3.6: Robustness Checks

(DV = Allocation in a Trader Dyad)

	Model 10	**Model 11**	**Model 12**	**Model 13**	**Model 14**
Learner Exp (Log)	1.227	1.580	0.0524	-0.391	0.937
	(1.004)	(1.036)	(1.085)	(1.165)	(1.208)
Guru Exp (Log)	5.859***	5.659***	5.658***	6.439***	6.767***
	(1.489)	(1.533)	(1.589)	(1.700)	(1.957)
Learner's Count of Gurus (Log)	0.0181	0.0844	-0.0206	0.0452	0.0861
	(0.0904)	(0.0915)	(0.100)	(0.105)	(0.114)
Count of Profitable Week (Log)	-1.282***	-1.369***	-1.432***	-1.079*	-0.384
	(0.353)	(0.371)	(0.399)	(0.423)	(0.459)
Count of Message Senders (Log)	-2.111***	-2.132***	-2.223***	-2.172***	-2.223***
	(0.173)	(0.180)	(0.193)	(0.204)	(0.214)
Count of Copiers	0.0198***	0.0198***	0.0207***	0.0197***	0.0202***
	(0.000685)	(0.000712)	(0.000750)	(0.000784)	(0.000903)
Gain in Pair	0.0199***	0.0196***	0.0189***	0.0200***	0.0222***
	(0.00121)	(0.00136)	(0.00140)	(0.00157)	(0.00178)
Geographic Distance	-0.973***	-0.979***	-0.728**	-1.004***	-0.726*
	(0.254)	(0.248)	(0.267)	(0.283)	(0.285)
Count of Copiers×Geographic Distance	0.000770**	0.000913***	0.000747**	0.000911**	0.000653*
	(0.000249)	(0.000257)	(0.000274)	(0.000283)	(0.000324)
Gain in Pair×Geographic Distance	-0.00230*	-0.00232*	-0.00262*	-0.00310**	-0.00298*
	(0.00113)	(0.00116)	(0.00119)	(0.00119)	(0.00145)
N	85290	75418	65358	55756	45622
Adj. R^2	0.0472	0.0474	0.0493	0.0482	0.0521
AIC	719528.3	635055.7	548940.4	466961.8	380679.9
BIC	719621.8	635148.0	549031.2	467051.0	380767.2
Learner Fixed Effects	Y	Y	Y	Y	Y
Guru Fixed Effects	Y	Y	Y	Y	Y
Observation Week Fixed Effects	Y	Y	Y	Y	Y

Robust standard errors clustered on Learners are reported in parentheses
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

6.5 Robustness check III: Are Guru's country-specific investments driving the allocation?

Finally, we test whether the effects are contingent on the Learner's preference for country-specific investment instruments. For instance, on XTrader, traders could invest in different stock exchange indices such as DAX, which include 30 German corporations listed on Frankfurt stock exchange and FTSE 100, which includes 100 UK firms listed on London Stock Exchange. Thus, it is possible that the allocation to a Guru may be subject to her country-specific investments, indicated by the frequency of her trades in the various stock exchange indices. To assess this concern, we first obtain a weekly count of transactions of stock indices for each Guru. Next, we aggregate this count over the entire observation window. Finally, we exclude all the dyads that include Gurus who have invested more than once in the stock exchange indices during the entire observation window. We argue that excluding such dyads should lead to substantially different results if the Guru's location-specific investments drive the allocation. However, as shown in Model 15, we obtain consistent results, indicating that Learner's preference for a Guru is not subject to the latter's country-specific investments (Table 3.7).

Table 3.7: Additional Robustness Checks

(DV = Allocation in a trader dyad)

	Model 15
Learner Exp (Log)	0.742
	(0.968)
Guru Exp (Log)	5.606***
	(1.431)
Learner's Count of Gurus (Log)	0.0107
	(0.0860)
Count of Profitable Week (Log)	-1.373***
	(0.339)
Count of Message Senders (Log)	-2.163***
	(0.171)
Count of Copiers	0.0200***
	(0.000651)
Gain in Pair	0.0192***
	(0.00116)
Geographic Distance	-0.942***
	(0.246)
Count of Copiers×Geographic Distance	0.000818***
	(0.000243)
Gain in Pair×Geographic Distance	-0.00208*
	(0.00103)
N	95227
Adj. R^2	0.0468
AIC	804666.4
BIC	804761.1
Learner Fixed Effects	Y
Guru Fixed Effects	Y
Observation Week Fixed Effects	Y

Robust standard errors clustered on Learners are reported in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7. CONCLUSION

The primary objective of our study is to explain the heterogeneity in the focal user's preferences for different information sources. To that end, our study provides several insights. Leveraging a highly granular and temporal data, we show that the focal user's preference for an information source associates positively with the source's herding signal (Shi & Whinston, 2013) as well as by the prior payoffs obtained from that source (Kaustia & Knüpfer, 2008). More importantly, however, we find that both these attributes are contingent on the physical distance between the information source and the focal member of the audience. In particular, the results indicate that as the physical distance between the two increases, the herding signal becomes more influential while prior payoffs are discounted. In this section, we discuss the implications of these findings for several streams of research.

First, our study contributes to the existing literature on the effect of physical distance on behavior in online platforms. Recent studies have shown that physical distance remains an impediment to interpersonal exchanges even in the age of the internet (Hortacsu, Martinez-Jerez, & Douglas, 2009). However, the majority of the empirical investigation in this regard has focused on the direct effect of distance on the outcome while the limited evidence pertaining to the interaction between the physical distance and other attributes of the source remains inconclusive (Agarwal, Catalini, & Goldfarb, 2015). We contribute to this literature by showing that the physical distance not only directly influences the focal user's preference for an information source but also gets factored into the assessment of other attributes of the information source. In sum, our study expands on the current understanding of how physical distance influences behavior in online platforms.

Specifically, our paper contributes to the literature on herding signal and physical distance in online platforms. Buiding on recent efforts (Forman et al., 2008l Agarwal et al., 2015), we show that the effect of herding signal is contingent on the physical distance. This finding is interesting because it calls into question the applicability of known explanations for such effect. For instance, Agarwal et al. (2015) argue that to the focal user, the herding signal about a physically proximate information source matters less because an observer may have prior connections with the proximal source (p. 271) enabling them to make their own determination of the quality of the artist's work. This is not the case with artists who are distant Thus, Agarwal et al. (2015) explain the interaction between herding signal and physical distance using the prior social ties between the proximal source and observer. However, we find that in settings, in which the presence of pre-existing social ties is unlikely, the herding signal about a physically proximate information source continues to matter less than that of the physically distant source. Thus, our finding shows that pre-existing social ties may not fully explain the interaction between the physical distance and herding signal, presenting an opportunity for future research to uncover other underlying theoretical mechanisms.

Next, our study also advances the understanding of how prior payoffs influence future actions. Extant literature suggests that prior payoffs may "reinforce choices" and hence, significantly influence future actions (Kaustia & Knüpfer, 2008). Consistent with this argument, we show that the prior payoffs from an information source have a positive effect on the focal user's subsequent preference towards that source. However, the physical distance between a focal user and the information source mitigates the effect of prior payoffs. This finding speaks directly to the competing explanations for the distance effect in online financial platforms (Lin & Viswanathan 2015). On the one hand, the *economic* explanation suggests that an individual prefers proximal investments because they provide better returns. On the other hand, the *behavioral* explanation indicates that factors such familiarity, emotional attachment, and identity come into play. The attenuating effect of physical distance on prior payoffs indicates that a behavioral explanation is driving the physical distance effect.

Finally, our study advances the literature on user behavior in the emerging "Fintech" phenomenon. In particular, we focus on social trading platforms (Doering, Neumann, & Paul, 2015). Although these platforms are nascent, with early entrants emerging in the first decade of the 21st century, they are already attracting the attention of network scholars, with an emphasis on copy trading (Wohlgemuth, Berger, & Wenzel, 2016; Pentland, 2013). In this regard, our study models the Learner's preference for a specific Guru's investment signals. Thus, our study is an apparent extension of recent exploratory efforts to study investment behavior in social trading (Glaser & Risius, 2017). Also, by showing theoretically consistent effects, which relate to multiple streams of extant literature, our study provides several conceptual anchors for future social trading researchers to examine the phenomenon from a behavioral standpoint. For instance, future work can assess whether these effects are transferable to different types of social trading platforms, and more generally, Fintech platforms.

Our study's findings must be viewed in light of certain limitations. We do not account for any cross-platform dynamics. It is possible that the same trader may be present on multiple social trading platforms and therefore, her behavior on XTrader may also be subject to her activity on the other platforms. We call for future research, incorporating cross-platform effects. Also, future work can also replicate our findings at different time intervals. Given the financial nature of the phenomenon, the user activity on most Fintech platforms is also driven by exogenous changes. While our study accounts for such changes by incorporating time-dummies, future work can replicate our study across different time windows to assess the temporal generalizability of our findings.

In closing, our study argues that while the literature on online platforms establishes the link between the ambient social information and users' behavior, little is understood about the heterogeneity in the focal user's preference for different sources of such information. Addressing this gap, we empirically and theoretically demonstrate the interplay between several underlying mechanisms. Going forward, we call for additional research examining determinants of information source preferences in online platforms, particularly in contexts in which such preferences may be costly.

CONCLUSION

While online platforms have made headways in several domains of human activity, none is perhaps of greater economic significance than the financial and investment markets (Lee & Shin, 2018). Such platforms are becoming increasingly mainstream as the millennials engage in financial investments (Morris, 2016). They are also interesting because they allow the participants to conduct a personal activity in a transparent and public fashion. This shift poses several interesting questions from a behavioral standpoint. Thus, it is hardly surprising that scholars have called for research specifically involving the online financial and investment platforms (Gomber, Kauffman, & Weber, 2015; Hendershott et al., 2017). In this stream, one of the important questions is *whether and how a user's investment behavior becomes susceptible to others' actions and decisions* especially in online platforms, which constantly expose users to information about "what others do"(Aral & Walker, 2011). The present dissertation addresses this question through a set of related empirical inquiries.

The dissertation makes several important contributions. First, using a highly granular dataset of person-level investments observed longitudinally, *we are able to empirically tease out the underlying theoretical mechanisms that are mostly latent in traditional settings.* In chapter 2, we show that in the presence of an audience, an individual becomes less susceptible to others' actions. Moreover, this effect is robust to exogenous changes in uncertainty. Assessing the "audience effect" as a driver of investment behavior is problematic in traditional contexts such as investment and securities firms (Gurun, Stoffman, & Yonker, 2018). Although the investment activity in these contexts may have a significant digital component, it is still personal in nature as the investment signals are rarely broadcast. Therefore, it is difficult to argue about as well as measure the audience of an investor in such settings. The behavioral implication of audience also speaks to the larger issue of increasing transparency of actions in digital space. We build on the ongoing effort to examine how "being visible online" changes the way individuals behave (Huang, Hong, & Burtch, 2016; Bapna et al., 2016).

Second, using the nuances of the empirical context, *we can unambiguously observe the extent to which an individual is susceptible to others' actions*. In the studies reviewed in chapter 1, we observe that the typically used behavioral outcomes include aggregated online purchases (Chevalier & Mayzlin, 2006), monetary donations (Burtch, Ghose, & Wattal, 2013), and adoption of digital goods and services (Aral & Walker, 2011; Bapna & Umyarov, 2015). However, these measures do not allow the researcher to attribute a focal user's behavior directly to others' actions. As a result, the researchers are limited in their ability to explain the focal user's susceptibility to information about others' actions. The novel outcome variables used in the two studies overcome this limitation. Thus, the present dissertation is an initial foray in isolating and explaining economic behavior in a highly social, transparent setting.

Lastly, *we resolve several theoretical tensions that have, in recent years, emerged as the economic activities are shifting to online settings*. For example, in chapter 2, we show that while a user who has independently attained success in the past is less susceptible to others in the present, the exogenous changes in uncertainty weaken this effect. Similarly, in chapter 3, we highlight the tension between prior payoffs that an information observer obtains from an information source and the physical distance between them. While higher payoffs in the past should increase the observer's preference for the source, physical distance should diminish it. We find that in the presence of these contrasting effects, the prior payoffs are significantly discounted. In the same vein, we show that the positive effect of herding signal about an information source attenuates the inhibiting effect of physical distance (Agarwal, Catalini, & Goldfarb, 2015; Greenwood & Gopal, 2017). In the closing section of each chapter, we have already discussed the research and practical implications of these findings in greater detail.

The dissertation opens up several directions for subsequent research. While we have described some of these in Chapters 2 and 3, we now suggest a broader research agenda pertaining to investment behavior in online platforms. First, future research can adopt a social network lens to understand online investment behavior. Given the availability of rich datasets on person-level economic activities, one can examine how interpersonal networks emerge and evolve in such settings. Some of the interesting research issues in this direction include the topology of interpersonal networks (Chen, Chen, & Xiao, 2013), and creation and persistence of ties (Dahlander & McFarland, 2013). Moreover, in at least some of the online investment and financial platforms, each tie that an investor forges is accompanied by a financial transaction. These platforms represent an important deviation in online networking space because, in more prevalent social networking platforms such as Facebook and LinkedIn, an individual incurs little financial cost in forming and maintaining network ties. Thus, scholars can also compare the individuals' decisions to form and maintain networks between the "costly" and relatively "costless" platforms to ascertain how social networking behavior changes when economic implications are introduced. Lastly, studies can examine the effects of prominent, offline events on individual behavior in digital space. We believe that given the direct linkage between investment activity and real-world events, online investment and financial platforms can act as a powerful setting to study the impact of offline incidences on online behavior.

BIBLIOGRAPHY

Agarwal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions. *Journal of Economics and Management Strategy, 24*(2), 253–274.

Agarwal, R., Sambamurthy, V., & Stair, R. (2000). The Evolving Relationship Between General and Specific Computer Self-Efficacy—An Empirical Assessment. *Information Systems Research, 11*(4), 418-430.

Aral, S. (2011). Commentary—Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science, 30*(2), 217-223.

Aral, S., & Walker, D. (2011). Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science, 57*(9), 1623-1639.

Aral, S., & Walker, D. (2012). Identifying Influential and Susceptible Members of Social Networks. *Science, 337*(6092), 337-341.

Aral, S., & Walker, D. (2014). Tie Strength, Embeddedness, and Social Influence: A Large-Scale, Networked Experiment. *Management Science, 60*(6), 1352-1370.

Archak, N., Ghose, A., & Ipeirotis, P. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science, 57*(8), 1485-1509.

Bandura, A. (1977). Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review, 84*, 191-215.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: Freeman.

Bapna, R., & Umyarov, A. (2015). Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks. *Management Science, 61*(8), 1902-1920.

Bapna, R., Ramaprasad, J., Shmueli, G., & Umyarov, A. (2016). One-Way Mirrors in Online Dating: A Randomized Field Experiment. *Management Science, 62*(11), 3100–3122.

Bapna, R., Gupta, A., Rice, S., & Sundararajan, A. (2017). Trust and The Strength of Ties in Online Social Networks: An Exploratory Field Experiment. *MIS Quarterly*, *41*(1), 115–130.

Baumeister, R. (1982). A Self-Presentational View of Social Phenomenon. *Psychological Bulletin, 91*, 3-26.

Beckmann, D., Menkhoff, L., & Suto, M. (2008). Does Culture Influence Asset Managers' Views and Behavior. *Journal of Economic Behavior & Organization, 67*(3-4), 624-643.

Berger, E., Wenzel, M., & Wohlgemuth, V. (2017). Imitation-Related Performance Outcomes in Social Trading: A Configurational Approach. *Journal of Business Research*. doi:https://doi.org/10.1016/j.jbusres.2017.12.016

Beugelsdijk, S., & Frijns, B. (2010). A Cultural Explanation of The Foreign Bias in International Asset Allocation. *Journal of Banking and Finance, 34*, 2121-2131.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy, 100*(5), 992-1026.

Bi, S., Liu, Z., & Usman, K. (2017). The Influence of Online Information on Investing Decisions of Reward-Based Crowdfunding. *Journal of Business Research*, *71*, 10–18.

Bond, C. (1982). Social Facilitation: A Self-Presentational View. *Journal of Personality and Social Psychology, 42*, 1042-1050.

Bond, C., & Titus, L. (1983). Social Facilitation: A Meta-Analysis of 241 Studies. *Psychological Bulletin, 94*(2), 265-292.

Borah, A., & Tellis, G. (2014). Make, Buy, or Ally? Choice of and Payoff from Announcements of Alternate Strategies for Innovation. *Marketing Science, 33*(1), 114-133.

Boudreau, K., Lacetera, N., & Lakhani, K. (2011). Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science, 57*(5), 843-863.

Boyd, N., & Vozikis, G. (1994). The Influence of Self-Efficacy on the Development of Entrepreneurial Intentions and Actions. *Entrepreneurship: Theory and Practice, 18*(4), 63-77.

Burtch, G., Ghose, A., & Wattal, S. (2013). An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets. *Information Systems Research, 24*(3), 499-519.

Burtch, G., Ghose, A., & Wattal, S. (2014). Cultural Differences and Geography as Determinants of Online Pro-Social Lending. *MIS Quarterly*, 773-794.

Burtch, G., Ghose, A., & Wattal, S. (2016). Secret Admirers: An Empirical Examination of Information Hiding and Contribution Dynamics in Online Crowdfunding. *Information Systems Research, 27*(3), 478–496.

Burtch, G., Hong, Y., Bapna, R., & Griskevicius, V. (2017). Stimulating Online Reviews by Combining Financial Incentives and Social Norms. *Management Science*. doi:https://doi.org/10.1287/mnsc.2016.2715

Campbell, J., & Cecez-Kecmanovic, D. (2011). Communicative Practices in an Online Financial Forum During Abnormal Stock Market Behavior. *Information and Management*, *48*(1), 37–52.

Chan-Lau, J., Liu, E., & Schmittmann, J. (2015). Equity Returns in the Banking Sector in the Wake of Great Recession and the European Sovereign Debt Crisis. *Journal of Financial Stability, 16*, 164-172.

Chen, X., Chen, Y., & Xiao, P. (2013). The Impact of Sampling and Network Topology on the Estimation of Social Intercorrelations. *Journal of Marketing Research, 50*(1), 95-110.

Chen, Y., Harper, F., Konstan, J., & Li, S. (2010). Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens. *American Economic Review, 100*(4), 1358-1398.

Chen, Y., Wang, Q., & Xie, J. (2011). Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational. *Journal of Marketing Research, 48*(2), 238-254.

Chen, Z., & Berger, J. (2016). How Content Acquisition Method Affects Word of Mouth. *Journal of Consumer Research*, *43*(1), 86–102.

Chevalier, J., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research, 43*(3), 345-354.

Chua, R., Roth, Y., & Lemoine, J. (2015). The Impact of Culture and Creativity: How Cultural Tightness and Cultural Distance Affect Global Innovation Crowdsourcing Work. *Administrative Science Quarterly, 60*(2), 189-227.

Correia, S. (2016, 3). A Feasible Estimator for Linear Models with Multi-Way Fixed Effects. Retrieved from http://scorreia.com/research/hdfe.pdf

Corrado, C., & Miller Jr., T. (2005). The Forecast Quality of CBOE Implied Volatility Indexes. *The Journal of Futures Markets, 25*(4), 339-373.

Croson, R., & Shang, J. (2008). The Impact of Downward Social Information on Contribution Decisions. *Experimental Economics*, *11*(3), 221–233.

Croson, R., & Shang, J. (2013). Limits of the Effect of Social Information on the Voluntary Provision of Public Goods: Evidence from Field Experiments. *Economic Inquiry*, *51*(1), 473–477.

Dahlander, L., & McFarland, D. (2013). Ties that Last: Tie Formation and Persistence in Research Collaborations over Time. *Administrative Science Quarterly*, *58*(1), 69–110

Døskeland, T., & Pedersen, L. J. T. (2016). Investing with Brain or Heart? A Field Experiment on Responsible Investment. *Management Science*, *62*(6), 1632–1644.

Doering, P., Neumann, S., & Paul, S. (2015, 7 7). *A Primer on Social Trading Networks – Institutional Aspects and Empirical Evidence.* Retrieved from Social Science Research Network: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2291421

Duan, W., Gu, B., & Whinston, A. (2009). Informational Cascades and Software Adoption on the Internet: An Empirical Investigation. *MIS Quarterly, 33*(1), 23-48.

Fast, N., Burris, E., & Bartel, C. (2014). Managing to Stay in the Dark: Managerial Self-Efficacy, Ego Defensiveness, and the Aversion to Employee Voice. *Academy of Management Journal, 57*(4), 1013-1034.

Feller, J., Gleasure, R., & Treacy, S. (2017). Information Sharing and User Behavior in Internet-Enabled Peer-To-Peer Lending Systems: An Empirical Study. *Journal of Information Technology, 32*(2), 127–146.

Floyd, K., Freling, R., Alhoqail, S., Cho, H., & Freling, T. (2014). How Online Product Reviews Affect Retail Sales: A Meta-analysis. *Journal of Retailing, 90*, 217–232.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291–313.

Frey, B., & Meier, S. (2004). Social Comparisons and Pro-Social Behavior: Testing "Conditional Cooperation" in a Field Experiment. *American Economic Review, 94*(5), 1717-1722.

Gamble, J., Brennen, M., & McAdam, R. (2017). A Rewarding Experience? Exploring How Crowdfunding is Affecting Music Industry Business Models. *Journal of Business Research, 70*, 25-36.

Garbarino, E., & Strahilevitz, M. (2004). Gender Differences in The Perceived Risk of Buying Online and The Effects of Receiving a Site Recommendation. *Journal of Business Research, 57*(7), 768-775.

Gibson, C. (1999). Do They Do What They Believe What They Can Do? Efficacy and Group Effectiveness Across Tasks and Cultures. *Academy of Management Journal, 42*(2), 138-152.

Gino, F., & Moore, D. (2007). The Effects of Task Difficulty on Use of Advice. *Journal of Behavioral Decision Making, 20*(1), 21-35.

Glaser, F., & Risius, M. (2017). Effects of Transparency: Analyzing Social Biases on Trader Performance in Social Trading. *Journal of Information Technology, 33*(1), 19–30.

Goel, S., & Goldstein, D. (2014). Predicting Individual Behavior with Social Networks. *Marketing Science, 33*(1), 82-93.

Gomber, B., Kauffman, R., & Weber, B. (2015). *Financial IS, Underlying Technologies, and The Fintech Revolution.* Retrieved 1 10, 2018, from Journal of Management Information Systems: http://www.jmis-web.org/cfps/JMIS_SI_Fintech.pdf

Greenwood, B., & Gopal, A. (2017). Ending the Mending Wall: Herding, Media Coverage, and Co-Location in IT Entrepreneurship. *MIS Quarterly, 41*(3), 989-1007.

Gu, B., Tang, Q., & Whinston, A. (2013). The Influence of Online Word-of-Mouth on Long Tail Formation. *Decision Support Systems, 56,* 474-481.

Guo, W., & Wohar, M. (2006). Identifying Regime Changes in Market Volatility. *Journal of Financial Research, 29*(1), 79-93.

Gurun, U., Stoffman, M., & Yonker, S. (2018). Trust Busting: The Effect of Fraud on Investor Behavior. *The Review of Financial Studies, 31*(4), 1341–1376.

Halter, G., & Dean, A. (1971). *Decisions under Uncertainty.* Cincinnati: South-Western Publishing.

Hartog, D., & Belschak, F. (2012). When Does Transformational Leadership Enhance Employee Proactive Behavior? The Role of Autonomy and Role Breadth Self-Efficacy. *Journal of Applied Psychology, 97*(1), 194-202.

Harvey, C., Stewart, D., & Ewing, M. (2011). Forward or Delete: What Drives Peer-to-Peer Message Propagation Across Social Networks? *Journal of Consumer Behavior, 10*, 365-372.

Hashim, M., Kannan, K., & Maximiano, S. (2017). Information Feedback, Targeting, and Coordination: An Experimental Study. *Information Systems Research, 28*(2), 289-308.

Hendershott, T., Zhang, M., Zhao, J., & Zheng, E. (2017). Call for Papers—Special Issue of Information Systems Research Fintech – Innovating the Financial Industry Through Emerging Information Technologies. *Information Systems Research, 28*(4), 885-886.

Hitsch, G., Hortaçsu, A., & Ariely, D. (2010). Matching and Sorting in Online Dating. *American Economic Review, 100*(1), 130-163.

Hong, H., Xu, D., Wang, G., & Fan, W. (2017). Understanding the Determinants of Online Review Helpfulness: A Meta-Analytic Investigation. *Decision Support System, 102,* 1–11.

Hortacsu, A., Martinez-Jerez, F., & Douglas, J. (2009). The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre. *American Economic Journal: Microeconomics, 1*(1), 53-74.

Horton, J. (2014). Misdirected Search Effort in a Matching Market: Causes, Consequences, and a Partial Solution. *Proceedings of The Fifteenth ACM Conference on Economics and Computation* (pp. 357-357). New York: ACM.

Huang, N., Hong, Y., & Burtch, G. (2016, 11 18). *Social Network Integration and User Content Generation: Evidence from Natural Experiments.* Retrieved from Social Science Research Network: https://ssrn.com/abstract=2872275

Huberman, G. (2001). Familiarity Breeds Investment. *Review of Financial Studies*, *14*(3), 659-680.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-292.

Kane, G., Alavi, M., Labianca, G., & Borgatti, S. (2014). What’s Different About Social Media Networks? A Framework and Research Agenda. *MIS Quarterly, 38*(1), 274-304.

Kang, L., Jiang, Q., & Tan, C. H. (2017). Remarkable Advocates: An Investigation of Geographic Distance and Social Capital for Crowdfunding. *Information and Management*, *54*(3), 336–348

Katona, Z., Zubcsek, P., & Sarvary, M. (2011). Network Effects and Personal Influences: The Diffusion of an Online Social Network. *Journal of Marketing Research, 48*(3), 425-443.

Kaustia, M., & Knüpfer, S. (2008). Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *Journal of Finance, 63*(6), 2679-2702.

Kim, J., & Ratchford, B. (2012). Consumer Choice and Use of Multiple Information Source for Automobile Purchases. *International Journal of Electronic Commerce, 16*(3), 7-40.

Koh, T., & Fichman, M. (2014). Multihoming Users' Preferences for Two-Sided Exchange Network. *MIS Quarterly, 38*(4), 977-996.

Kondor, D., Pósfai, M., Csabai, I., & Vattay, G. (2014). Do the Rich Get Richer? An Empirical Analysis of the Bitcoin Transaction Network. *Plos One, 9*(2).

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. (2016). From Social to Sale: The Effects of Firm-Generated Content in Social Media on Customer Behavior. *Journal of Marketing, 80*, 7–25.

Lee, F. (2002). The Social Costs of Seeking Help. *Journal of Applied Behavioral Science, 38*(1), 17-35.

Lee, C. J., & Andrade, E. B. (2011). Fear, Social Projection, and Financial Decision Making. *Journal of Marketing Research*, *48*, S121–S129.

Lee, I., & Shin, Y. (2018). Fintech: Ecosystem, Business Models, Investment Decisions, and Challenges. *Business Horizons, 61*, 35-46.

Leonardi, P. (2015). Ambient Awareness and Knowledge Acquisition: Using Social Media to Learn "Who Knows What" and "Who Knows Whom." *MIS Quarterly, 39*(4), 747-762.

Lin, M., & Viswanathan, S. (2015). Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market. *Management Science, 62*(5), 1393-1414.

Lin, M., Lucas, H., & Shmueli, G. (2013). Too Big to Fail: Large Samples and the p-Value Problem. *Information Systems Research, 24*(4), 906-917.

Liu, D., Brass, D., Lu, Y., & Chen, D. (2015). Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding. *MIS Quarterly, 39*(3), 729-742.

Liu, S., Wang, M., Liao, H., & Shi, J. (2014). Self-Regulation During Job Search: The Opposing Effects of Employment Self-Efficacy and Job Search Behavior Self-Efficacy. *Journal of Applied Psychology, 99*(6), 1159-1172.

Liu, Z., & Park, S. (2015). What Makes a Useful Online Review? The Implication for Travel Product. *Tourism Management, 45*, 140-151.

Ludwig, S., Ruyter, K., Friedman, M., Brüggen, E., Wetzels, M., & Pfann, G. (2013). More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing, 77*, 87 – 103.

Lukkarinen, A., Teich, J. E., Wallenius, H., & Wallenius, J. (2016). Success Drivers of Online Equity Crowdfunding Campaigns. *Decision Support Systems*, *87*, 26–38.

MacNab, B., & Worthley, R. (2007). Self-Efficacy as an Interpersonal Predictor of Internal Whistleblowing: A US and Canada Examination. *Journal of Business Ethics, 79*, 407-421.

Mayer, T., & Zignago, S. (2011, 12). *Notes pm CEPII's Distance Measures: The GeoDist Database.* Retrieved from CEPII: http://www.cepii.fr/PDF_PUB/wp/2011/wp2011-25.pdf

McFarland, L., & Ployhart, R. (2015). Social Media: A Contextual Framework to Guide Research and Practice. *Journal of Applied Psychology, 100*(6), 1653–1677.

Morris, E. (2016, 12). *Fintech Analyst Report Part-3: Asset Management.* Retrieved from Pitchbook: https://files.pitchbook.com/pdf/PitchBook_Fintech_Analyst_Report_Part_3_Asset_Management.pdf

Newberry, P. (2016). An Empirical Study of Observational Learning. *The RAND Journal of Economics*, *47*(2), 394–432.

Oleson, K., Poehlmann, K., Yost, J., Lynch, M., & Arkin, R. (2000). Subjective Overachievement: Individual Differences in Self-Doubt and Concern With Performance. *Journal of Personality*, *68*(3), 491–524.

Papke, L., & Wooldridge, J. (2008). Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates. *Journal of Econometrics, 145*(1-2), 121-133.

Pentland, A. (2013). Beyond the Echo Chamber. *Harvard Business Review, 91*(11), 80-86.

Pfattheicher, S., & Keller, J. (2015). The Watching Eyes Phenomenon: The Role of A Sense of being Seen and Public Self-Awareness. *European Journal of Social Psychology, 45*(5), 560–566.

Rai, A., Arikan, I., Pye, J., & Tiwana, A. (2015). Fit and Misfit of Plural Sourcing Strategies and IT-Enabled Process Integration Capabilities: Consequences of Firm Performance in the U.S. Electric Utility Industry. *MIS Quarterly, 39*(4), 865-885.

Rao, H., Greve, H., & Davis, G. (2001). Fool's Gold: Social Proof in the Initiation and Abandonment of Coverage by Wall Street Analysts. *Administrative Science Quarterly, 46*(3), 502-526.

Ray, S., Kim, S., & Morris, J. (2014). The Central Role of Engagement in Online Communities. *Information Systems Research, 25*(3), 528 - 546.

Reingen, P. (1982). Test of a List Procedure for Inducing Compliance with a Request to Donate Money. *Journal of Applied Psychology*, *67*(1), 110-118.

Rhue, L., & Sundararajan, A. (2013). How Digital Social Visibility Shapes Consumer Choices. Retrieved from Social Science Research Network: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2347793

Richter, A., Hirst, G., Knippenberg, D., & Baer, M. (2012). Creative Self-Efficacy and Individual Creativity in Team Contexts: Cross-Level Interactions With Team Informational Resources. *Journal of Applied Psychology, 97*(6), 1282–1290.

Roma, P., Petruzzelli, A., & Perrone, G. (2017). From the Crowd to the Market: The Role of Reward-Based Crowdfunding Performance in Attracting Professional Investors. *Research Policy*, *46*(9), 1606–1628.

Rosario, A., Sotgiu, F., Valck, K., & Bijmolt, T. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research*, *53*(3), 297–318.

Sarwar, G. (2012). Intertemporal Relations between The Market Volatility Index and Stock Index Returns. *Applied Financial Economics, 22*(11), 899-909.

Sas, C., & Khairuddin, I. (2017). Design for Trust: An Exploration of the Challenges and Opportunities of Bitcoin Users. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6499-6510). Denver: ACM.

Shang, J., & Croson, R. (2009). A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *The Economic Journal, 119*(540), 1422-1439.

Shang, J., Reed, A., & Croson, R. (2008). Identity Congruence Effects on Donations. *Journal of Marketing Research, 45*(3), 351-361.

Shea, C., Davisson, E., & Fitzsimons, G. (2013). Riding Other People's Coattails: Individuals with Low Self-Control Value Self-Control in Other People. *Psychological Science, 24*(6), 1031-1036.

Shi, Z., & Whinston, A. (2013). Network Structure and Observational Learning: Evidence from a Location-Based Social Network. *Journal of Management Information Systems, 30*(2), 185–212.

Sul, H., Dennis, A., & Yuan, L. (2014). Trading on Twitter: The Financial Information Content of Emotion in Social Media. *47th Hawaii International Conference on System Sciences* (pp. 806-815). Waikoloa: IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/6758703/

Sun, H. (2013). A Longitudinal Study of Herd Behavior in the Adoption and Continued Use of Technology. *MIS Quarterly, 37*(4), 1013-1041.

Sundén, A., & Surette, B. (1998). Gender Differences in the Allocation of Assets in Retirement Savings Plans. *American Economic Review, 88*(2), 207-211.

Thatcher, J., & Perrewe, P. (2002). An Empirical Examination of Individual Traits as Antecedents to Computer Anxiety and Computer Self-Efficacy. *MIS Quarterly, 26*(4), 381-396.

Thies, F., Wessel, M., & Benlian, A. (2016). Effects of Social Interaction Dynamics on Platforms. *Journal of Management Information Systems, 33*(3), 843-873.

Tierney, P., & Farmer, S. (2002). Creative Self-Efficacy: Its Potential Antecedents and Relationship to Creative Performance. *Academy of Management Journal, 45*(6), 1137-1148.

Torres, S., Portugal, P., Addison, J., & Guimarães, P. (2013, 9). The Source of Wage Variation: A Three-Way High Dimensional Fixed Effects Regression Model. Retrieved from Bank of Portugal: https://www.bportugal.pt/sites/default/files/anexos/papers/wp201309.pdf

Walker, W., Harremoes, P., van der Sluijs, J., van Asselt, M., Janssen, P., & von Krauss, M. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessments, 4*(1), 5-17.

Wang, Y., Rod, M., Ji, S., & Deng, Q. (2017). Social Media Capability in B2B Marketing: Toward a Definition and a Research Model. *Journal of Business & Industrial Marketing*, *32*(8), 1125-1135.

Wessel, M., Thies, F., & Benlian, A. (2016). The Emergence and Effects of Fake Social Information: Evidence from Crowdfunding. *Decision Support Systems*, *90*, 75–85.

Willemsen, L., Neijens, P., Bronner, F., & de Ridder, J. (2011). ''Highly Recommended!'' The Content Characteristics and Perceived Usefulness of Online Consumer Reviews. *Journal of Computer-Mediated Communication*, 19–38.

Wohlgemuth, V., Berger, E., & Wenzel, M. (2016). More than Just Financial Performance: Trusting Investors in Social Trading. *Journal of Business Research*, *69*(11), 4970–4974.

Wu, L. (2013). Social Network Effects on Productivity and Job Security: Evidence from the Adoption of Social Networking Tool. *Information Systems Research*, *24*(1), 30-51.

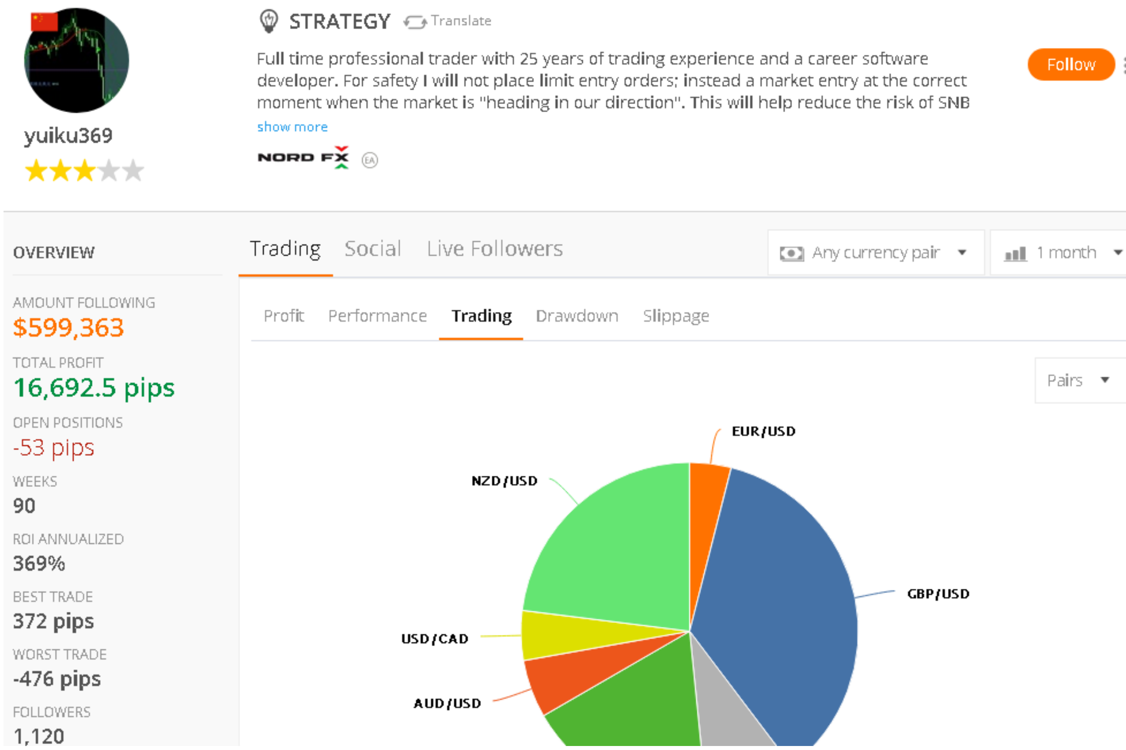
Yao, Y., Dresner, M., & Palmer, J. (2009). Private Network EDI vs. Internet Electronic Markets: A Direct Comparison of Fulfillment Performance. *Management Science, 55*(5), 843-852.

Yeo, G., & Neal, A. (2013). Revisiting the Functional Properties of Self Efficacy: A Dynamic Perspective. *Journal of Management*, *39*(6), 1385-1396.

Yin, D., Bond, S., & Zhang, H. (2014). Anxious or Anger? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews. *MIS Quarterly*, *38*(2), 539-560.

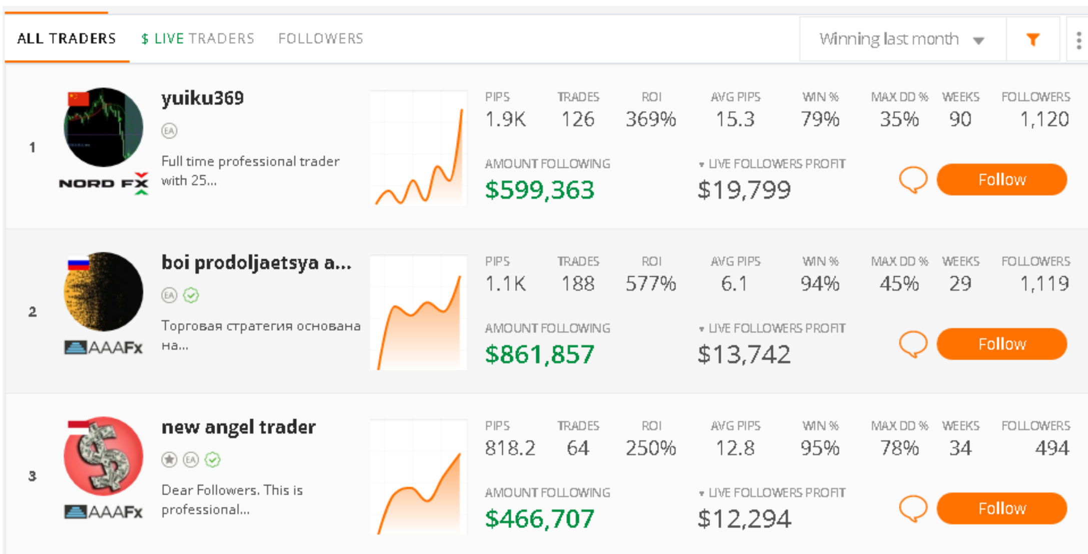
APPENDICES

Appendix I: Social information capturing the trader's investment activity



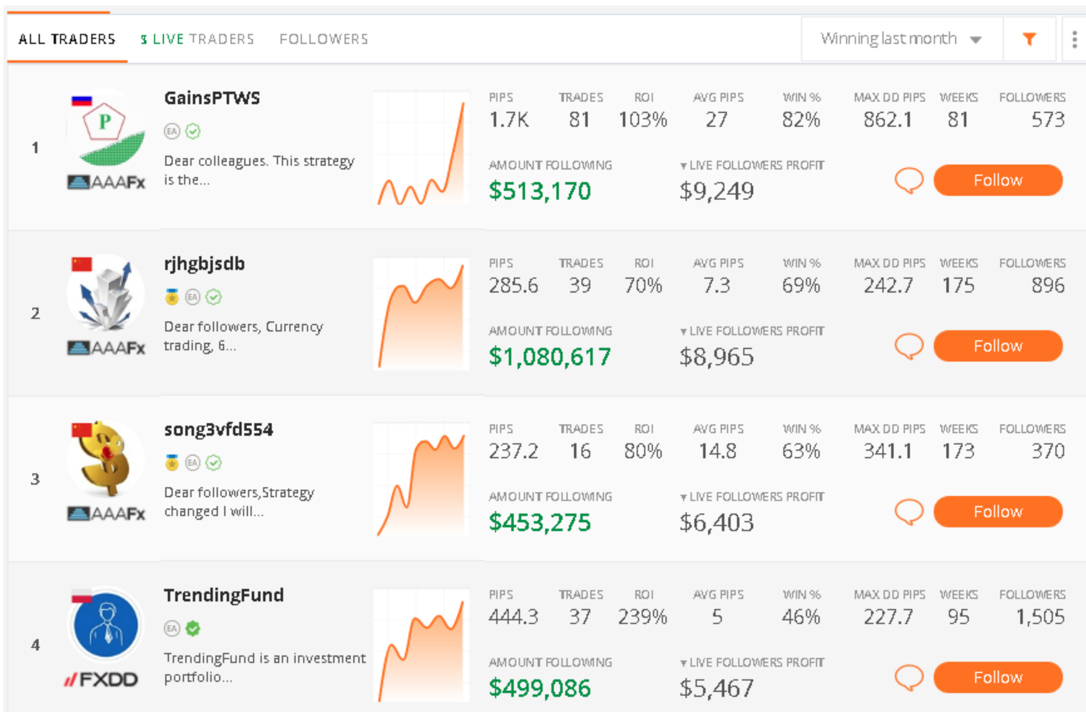
As an illustration, we show a trader profile page from a prominent social trading platform named ZuluTrade. The platform's design features are similar to XTrader. The profile is publicly visible at https://www.zulutrade.com/trader/318058/trading?t=30 and was accessed on 22nd October 2017. Note that the platform shows a host of actions a trader undertook, resulting in a trove of ambient social information. In the right top corner, we see a button named as "Follow." Any trader can click on this button to rely on this particular trader's actions. On other platforms, including XTrader, the profile pages of the traders are shown similarly.

Social trading platforms also provide an aggregate view of the ambient social information. In the example given below, we see that the actions of several traders are visible simultaneously along with the “Follow” button. This dashboard is publicly available on https://www.zulutrade.com/traders and was accessed on 22nd October 2017. The two images indicate that ambient social information, comprising the trading activities of other traders, as well as the feature for copy trading, are readily visible to the users on social trading platforms. Collectively, these features enable users to rely on ambient social information.



Appendix II: How a focal user (Learner) can search for an information source (Guru)?

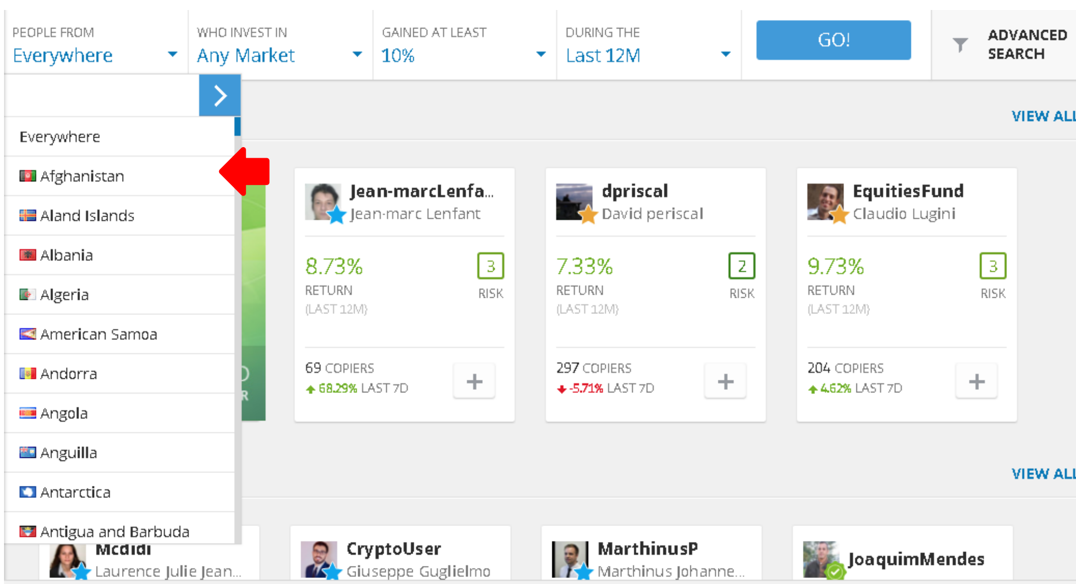
Social trading platforms, including XTrader, provide several ways for the Learner to find other traders for copy-trading. The illustration below shows an aggregation of active traders on a leading social trading platform called ZuluTrade (https://www.zulutrade.com/traders)



In addition, platforms provide country-based trader search, which underscores the importance of physical distance. An illustration below shows the country-based trader search for ZuluTrade (Figure A: http://v4.zulutrade.com/PerformanceProvidersByCountry.aspx) and eToro (Figure B: https://www.etoro.com/discover/people)

Figure A

Figure B



Appendix III: Comparing the social trading platforms

The hypothesized relationships in our study depend on the availability of the relevant information to the traders. For instance, in a trader dyad, estimating the effect of expertise signal on allocation is reasonable if the traders have access to such information. Moreover, to assess generalizability, it is also essential to demonstrate that the availability of such information cues transcends the platform used as the empirical context. In this appendix, we compare the available information across several leading social trading platforms: eToro (https://www.etoro.com/), ZuluTrade (https://www.zulutrade.com/), Tradeo (https://tradeo.com/), and Ayondo (https://www.ayondo.com/). We find that the availability of the crucial information is reasonably consistent across the platforms.

Information on given trader's profile page	**eToro**	**ZuluTrade**	**Tradeo**	**Ayondo**
Number of copying traders (Visible prominently on each trader's profile)	✓	✓	✓	✓
Prior payoffs from the particular information source (Accessible through each trader's portfolio)	✓	✓	✓	✓
Home country (Shown typically close to the trader's username)	✓	✓	✓	✗

Appendix IV: Summary of Fintech Studies

(These studies examine the effect of information about others' actions on the observer's behavior)

Paper	**Empirical Context**	**Mapping with Classification Framework (Chapter 1)**
Thies, Wessel & Benlian (2016)	IndieGogo.com, Facebook	Presentation
Bi, Liu & Usman (2017)	Zhongchou.com[A Chinese Crowdfunding website]	
Campbell & Cecez-Kecmanovic (2011)	HotCopper (Internet Discussion Forum)	Content
Hashim, Kannan & Maximiano (2011)	Lab experiment	
Burtch, Ghose & Wattal (2013)	A crowdfunding platform for Online Journalism	
Døskeland & Pedersen (2016)	A Norwegian online bank	
Thies, Wessel & Benlian (2016)*	IndieGogo.com, Facebook	
Wessel, Thies & Benlian (2016)	Kickstarter.com, Facebook.com	
Wohlgemuth, Berger & Wenzel (2016)	EToro [A social trading platform]	
Bi, Liu & Usman (2017)*	Zhongchou.com[A Chinese Crowdfunding website]	
Feller, Gleasure & Treacy (2017)	LendingClub [A P2P lending platform]	
Roma, Petruzzelli & Perrone (2017)	Technology projects on Kickstarter	
Lee & Andrade (2011)	Lab experiment	Observer and Source
Burtch, Ghose & Wattal (2016)	A large Crowdfunding platform	
Lukkarinen, et al. (2016)	Invesdor.com [A Crowdfunding platform]	
Bapna, et al. (2017)	Customized FB Application for Monetary Transfer	
Kang, Jiang & Tan (2017)	Demohour.com [A Chinese Crowdfunding platform) Weibo.com	

* Mapped on to multiple dimensions